



PCT

ORGANISATION MONDIALE DE LA PROPRIÉTÉ INTELLECTUELLE
Bureau international

DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITE DE COOPERATION EN MATIÈRE DE BREVETS (PCT)

<p>(51) Classification internationale des brevets ⁵ : C12N 15/48, 15/62, 15/86 C12N 5/10, A61K 39/21</p>	<p>A1</p>	<p>(11) Numéro de publication internationale: WO 92/19742 (43) Date de publication internationale: 12 novembre 1992 (12.11.92)</p>
<p>(21) Numéro de la demande internationale: PCT/FR92/00394 (22) Date de dépôt international: 30 avril 1992 (30.04.92) (30) Données relatives à la priorité: 91/05392 2 mai 1991 (02.05.91) FR (71) Déposant (pour tous les Etats désignés sauf US): TRANS- GENE S.A. [FR/FR]; 11, rue de Molsheim, F-67000 Strasbourg (FR). (72) Inventeur; et (75) Inventeur/Déposant (US seulement) : KIENY, Marie-Paule [FR/FR]; 7, rue Alois-Quintenz, F-67000 Strasbourg (FR). (74) Mandataire: SCHRIMPF, Robert; Cabinet Regimbeau, 26, avenue Kléber, F-75116 Paris (FR).</p>		<p>(81) Etats désignés: AT (brevet européen), BE (brevet euro- péen), CA, CH (brevet européen), DE (brevet européen), DK (brevet européen), ES (brevet européen), FR (brevet européen), GB (brevet européen), GR (brevet européen), IT (brevet européen), JP, LU (brevet européen), MC (brevet européen), NL (brevet européen), SE (brevet européen), US. Publiée <i>Avec rapport de recherche internationale. Avant l'expiration du délai prévu pour la modification des revendications, sera republiée si de telles modifications sont reçues.</i></p>
<p>(54) Title: NOVEL HYBRID, SOLUBLE AND UNCLEAVED gp160 VARIANT (54) Titre: NOUVEAU VARIANT gp160 NON-CLIVABLE, SOLUBLE, DE FORME HYBRIDE (57) Abstract A novel method for building an expression cassette comprising a DNA unit which codes for a precursor of a hybrid, unres- tricted and soluble gp160 variant, and the novel variant thereby obtained, are disclosed. (57) Abrégé L'invention se rapporte en particulier à une nouvelle méthode pour construire une cassette d'expression comportant une unité d'ADN codant pour un précurseur d'un variant gp160 hybride, non-clivable et soluble, ainsi qu'au nouveau variant men- tionné ci-dessus.</p>		

UNIQUEMENT A TITRE D'INFORMATION

Codes utilisés pour identifier les Etats parties au PCT, sur les pages de couverture des brochures publiant des demandes internationales en vertu du PCT.

AT	Autriche	FI	Finlande	ML	Mali
AU	Australie	FR	France	MN	Mongolie
BB	Barbade	GA	Gabon	MR	Mauritanie
BE	Belgique	GB	Royaume-Uni	MW	Malawi
BF	Burkina Faso	GN	Guinée	NL	Pays-Bas
BG	Bulgarie	GR	Grèce	NO	Norvège
BJ	Bénin	HU	Hongrie	PL	Pologne
BR	Brésil	IE	Irlande	RO	Roumanie
CA	Canada	IT	Italie	RU	Fédération de Russie
CF	République Centrafricaine	JP	Japon	SD	Soudan
CG	Congo	KP	République populaire démocratique de Corée	SE	Suède
CH	Suisse	KR	République de Corée	SN	Sénégal
CI	Côte d'Ivoire	LI	Liechtenstein	SU	Union soviétique
CM	Cameroun	LK	Sri Lanka	TD	Tchad
CS	Tchécoslovaquie	LU	Luxembourg	TG	Togo
DE	Allemagne	MC	Monaco	US	Etats-Unis d'Amérique
DK	Danemark	MG	Madagascar		
ES	Espagne				

5

"Nouveau variant gp160 non-clivable, soluble, de forme hybride"

10

15 La présente invention concerne une méthode pour construire des moyens nécessaires à la production de nouvelles molécules capables d'induire une réponse immuno-protectrice à l'encontre d'un virus responsable du syndrome d'immunodéficience acquise (SIDA).

Chez un individu, cette maladie se développe à la suite d'une infection des
20 lymphocytes T-helper par un rétrovirus HIV (human immunodeficiency virus). Jusqu'à présent on a répertorié ces rétrovirus en deux grand types: le type HIV-1 qui sévit essentiellement en Europe et en Amérique du Nord et le type HIV-2 caractéristique des infections africaines. A l'intérieur d'un même type, les rétrovirus HIV présentent entre eux un certain degré de variabilité qui se caractérise par exemple par une différence de
25 trophisme cellulaire ou par des protéines virales légèrement différentes les uns des autres. C'est pourquoi, lorsqu'on souhaite se référer à un rétrovirus HIV particulier, on emploie de préférence le terme de "souche virale".

D'une manière générale, les rétrovirus HIV ont la structure suivante : la molécule
30 d'ARN génomique et diverses protéines associées sont encapsulés dans une capside de nature protéique (nucléocapside). L'ensemble est protégé par une membrane d'origine cellulaire ayant incorporé la protéine d'enveloppe d'origine virale.

Cette protéine d'enveloppe, sous des formes diverses, est à ce jour considéré comme
35 un élément thérapeutique potentiel d'un vaccin contre le Sida.

Dans les conditions naturelles, la protéine d'enveloppe (env) est initialement synthétisée sous forme d'un précurseur comportant à son extrémité N-terminale une séquence signal qui initie le passage du précurseur dans le réticulum endoplasmique (voie
40 de sécrétion). Ce peptide signal est ensuite éliminé par clivage protéolytique. Le produit

- 2 -

de ce clivage est une protéine appelée gp160 qui est elle-même par la suite clivée en une petite sous-unité gp40 et une grande sous-unité gp120. L'extrémité N-terminale de la gp120 correspond à l'extrémité N-terminale de la gp160 tandis que l'extrémité C-terminale de la gp40 correspond à l'extrémité C-terminale de la gp160.

5

Chacune des sous-unités est sécrétée à l'extérieur de la cellule infectée. Toutefois, la gp40 reste ancrée dans la membrane cellulaire par son domaine transmembranaire. Sa partie C-terminale (domaine intra-cytoplasmique) reste en contact avec le cytoplasme tandis que sa partie N-terminale (domaine extra-cytoplasmique) se trouve à la surface de la cellule. La sous-unité gp120 est relarguée à l'extérieur de la cellule où elle interagit avec le domaine extracellulaire de la sous-unité gp40. Les deux sous-unités de la protéine d'enveloppe restent ainsi associées sous forme de complexe.

La séquence en acides aminés des précurseurs des protéines env de diverses souches virales sont maintenant connues. A titre d'exemple, la Figure 1 présente la séquence des précurseurs des gp160 des souches virales HIV-1 Bru, HIV-1 MN, HIV-1 ELI, HIV-1 RF, HIV-1 SF2C et HIV-1 SC. De même, la Figure 2 présente la séquence du précurseur de la gp160 de la souche virale HIV-2 Rod.

Par la suite et pour faciliter la compréhension, les séquences des gp160 autres que la gp160 de la souche HIV-1 Bru seront ci-après décrites par référence à la séquence de la gp160 de la souche HIV-1 Bru (ci-après appelée gp160-Bru) comme suit :

a) La gp160-Bru possède 831 acides aminés, ceux-ci numérotés de 1 à 831. En outre la gp120-Bru correspond à la séquence des 486 premiers acides aminés de la gp160-Bru tandis que la gp40-Bru correspond à la séquence commençant avec l'acide aminé en position 487 et finissant avec l'acide aminé en position 831.

b) La séquence d'une gp160 quelconque est alignée avec la séquence de la gp160-Bru de manière à présenter un maximum d'homologie. A cette fin, des emplacements vacants peuvent être introduits soit dans la séquence de la gp160-Bru, soit dans la séquence de la gp160 quelconque. Par définition, la position d'un acide aminé dans la gp160 quelconque sera spécifiée à l'identique par la position de l'acide aminé homologue dans la gp160-Bru.

35

c) En conséquence, lorsqu'un emplacement vacant est introduit dans la séquence de la gp160 quelconque, en face de l'acide aminé en position x, il n'existe pas d'acide aminé en position x dans la séquence de la gp160. Cependant, on considère que cet acide aminé jouit d'une présence virtuelle.

40

- 3 -

- d) Lorsqu'un ou plusieurs emplacement(s) vacant(s) est (sont) introduit(s) dans la séquence de la gp160-Bru entre l'acide aminé en position x et l'acide aminé en position $x + 1$, on considère que, dans la gp160 quelconque, la position recouvre au moins deux sous-positions x_a, x_b, \dots, x_n , chacune occupée par un acide aminé. Par définition, la position d'aval x_n représente la position x .
- 5
- e) Lorsqu'un ou plusieurs emplacement(s) vacant(s) est (sont) introduit(s) dans la séquence de la gp160-Bru en amont de l'acide aminé en position 1, on considère que, dans la gp160 quelconque, la position 1 recouvre au moins deux sous-positions $1_a, 1_b, \dots, 1_n$, chacune occupée par un acide aminé. Par définition, la position d'amont NH_2 -terminale 1_a représente la position 1.
- 10
- f) Lorsqu'un ou plusieurs emplacement(s) vacant(s) est (sont) introduit(s) dans la séquence de la gp160 quelconque en face du ou des acide(s) aminé(s) NH_2 -terminal(aux) de la gp160-Bru, on indique que, par définition, l'acide aminé en position NH_2 -terminale dans la gp160 quelconque cumule la position 1 et sa position d'acide aminé homologue.
- 15
- g) Lorsqu'un ou plusieurs emplacement(s) vacant(s) est (sont) introduit(s) dans la séquence de la gp160 quelconque en face du ou des acide(s) aminé(s) $COOH$ -terminal(aux) de la gp160-Bru, on indique que, par définition, l'acide aminé en position $COOH$ -terminale dans la gp160 quelconque cumule la position 841 et sa position d'acide aminé homologue.
- 20
- 25 A ce jour, dans l'optique d'un vaccin, la protéine d'enveloppe est considérée comme un candidat potentiellement intéressant d'un point de vue thérapeutique. Toutefois, sa structure multichaîne d'origine constitue un handicap en terme de faisabilité. C'est pourquoi, il est apparu préférable d'utiliser une protéine simple chaîne qui conserverait la plupart des épitopes de la gp120 et ceux de la gp40. Ce type de protéine a déjà été
- 30 proposé dans la demande de brevet WO 87/6260. Il s'agit plus particulièrement d'un variant gp160 non-clivable et soluble.

Toutes les gp160 des souches virales de type HIV-1 possèdent, en positions 483 - 486, un site de clivage dit majeur, reconnu par une ou des enzymes protéolytiques. Ce site

35 de clivage est le même pour toutes les gp160 décrites à la Figure 1 et correspond à la séquence Arg-Glu-Lys-Arg (REKR). Les enzymes protéolytiques coupent immédiatement en aval de ce site de clivage pour libérer une gp120 et une gp40.

Lorsque ce site de clivage n'existe pas, on a démontré que les enzymes

40 protéolytiques reconnaissent avec une efficacité toutefois moindre un site de clivage dit

- 4 -

mineur et d'effectuer une coupure immédiatement en aval de celui-ci. Là aussi, ce site de clivage est le même pour toutes les gp160 montrées à la Figure 1. Il correspond à la séquence Lys-Arg-Arg (KRR) en positions 477 - 479, selon certains auteurs ou selon d'autres, à la séquence Lys-Ala-Lys-Arg (KAKR) en positions 475 - 478.

5

En outre, on indique que les gp160 des souches virales de type HIV-2 possèdent uniquement un site de clivage majeur en positions 475 - 478 qui correspond à la séquence Lys-Glu-Arg-Lys (KEKR).

- 10 Un variant gp160 non-clivable est un analogue artificiel d'une gp160 native. Sa séquence en acides aminés correspond à celle d'une gp160 native dans laquelle le site de clivage majeur et/ou le site de clivage mineur a (ont) été supprimé(s).

- 15 Une tel variant gp160 peut être synthétisé grâce aux techniques de l'ADN recombinant. Il suffit d'isoler un fragment d'ADN codant pour une gp160 native puis de modifier la région codant pour le site de clivage majeur par mutagenèse dirigée, de manière à obtenir un fragment d'ADN codant pour un variant gp160 insensible à une action protéolytique. Par la suite ce dernier fragment d'ADN est exprimé dans des conditions appropriées pour donner ledit variant gp160. Un tel variant gp160 qui ne
20 contient plus le site de clivage majeur est appelée ci-après un variant gp160 non-clivable de type A.

- De manière préférée, la région codant pour le site de clivage mineur peut être en outre modifiée à des fins identiques. Dans ces conditions, on obtient un variant gp160
25 non-clivable de type A' qui ne contient ni de site le clivage majeur, ni le site de clivage mineur.

- A cette fin, on indique à titre d'exemple que le site de clivage REKR, KEKR ou KAKR et le site de clivage KRR peuvent être respectivement remplacés par les séquences
30 Asn-Glu-His-Gln (NEHQ) et Gln-Asn-His (QNH).

- D'autre part, il est aussi nécessaire qu'une gp160 soit obtenue sous forme soluble. Une telle gp160 correspond à une gp160 native dans laquelle la région transmembranaire de nature hydrophobe a été supprimée. Cette région transmembranaire est située dans la
35 zone correspondant à la gp40, du résidu acide aminé en position 659 au résidu acide aminé en position 680 De manière additionnelle mais superflue, une autre région hydrophobe, allant du résidu acide aminé en position 487 au résidu acide aminé en position 514 pourrait être déléetée.

- 40 De façon similaire, les moyens nécessaires à la synthèse d'un variant gp160 soluble

- 5 -

incluent bien évidemment le fragment d'ADN correspondant qui doit être obtenu par modification du fragment d'ADN original.

La comparaison des séquences des différentes gp160 déjà connues a mis en évidence
5 au moins trois domaines dont la séquence est hypervariable d'une gp160 à l'autre. Ces trois domaines sont couramment appelés les domaines (ou boucles) V_1 , V_2 et V_3 .

Les deux premiers domaines V_1 et V_2 sont situés entre le résidu cystéine en position 96 et le résidu cystéine en position 171 tandis que le troisième domaine V_3 est situé du
10 résidu cystéine en position 271 au résidu cystéine en position 306.

Il existe aussi un dernier domaine présentant un certain degré de variabilité, toutefois considéré comme moindre. Il s'agit du site de fixation au récepteur CD4 des lymphocytes T-helper; celui-ci étant situé approximativement du résidu acide aminé en position 340 au
15 résidu acide aminé en position 440.

Quelque soit la gp160 considérée, des expériences de vaccination ont montré que le troisième domaine hypervariable est essentiel pour obtenir une immunité convenable. Toutefois, en raison de la nature hypervariable de ce domaine, la protection développée
20 ne serait efficace qu'à l'encontre de la souche virale dont est issue la gp160 utilisée.

Enfin il semblerait que le premier et le deuxième domaines hypervariables ainsi que le site de fixation au récepteur CD4 influent sur le degré d'immunité que l'on pourrait obtenir.

25

Un vaccin d'un intérêt général devrait permettre de protéger les individus contre la majorité des souches virales HIV qui sévissent dans le monde. En conséquence, un vaccin à base de gp160 devrait contenir diverses gp160, chacune dérivée d'une souche virale différente. Afin de mettre en oeuvre un tel vaccin, il convient tout d'abord de construire
30 les moyens destinés à la production du variant gp160 non-clivable et soluble, correspondant à chaque souche virale. En première approche, il s'agirait donc à chaque fois de cloner le fragment d'ADN codant pour la gp160 d'une souche virale déterminée, d'en déterminer sa séquence, de modifier ce fragment d'ADN afin de supprimer par substitution ou délétion des sites de clivage et des régions hydrophobes, puis enfin de
35 placer ce fragment d'ADN ainsi modifié dans des conditions d'expression appropriées. Comme ce type d'opération devrait être renouvelée pour chaque gp160, la préparation d'un tel vaccin est prévue longue et coûteuse.

De plus, les souches virales peuvent varier au cours du temps e.g. par mutation.
40 Telle souche, cause majeure de l'infection dans telle région du globe à tel moment, peut

- 6 -

régresser en tant qu'agent infectieux et telle autre souche apparaître en remplacement. Il est donc important de pouvoir adapter rapidement un vaccin aux conditions épidémiologiques et, dans le cas présent, de disposer d'un variant gp160 dans les meilleurs délais.

5

A cette fin, on a maintenant trouvé une nouvelle méthode pour obtenir un variant gp160 approprié quelque soit la souche virale considérée. Dans la mesure où l'on dispose d'une cassette d'expression destinée à la production d'un variant gp160 non-clivable et soluble dérivé d'une première souche virale, cette méthode permet d'obtenir un variant
10 similaire dérivé d'une deuxième souche virale en évitant le clonage complet du fragment d'ADN codant pour la gp160 de la deuxième souche virale et les modifications qui devraient s'en suivre.

En conséquence, l'invention propose une méthode pour construire une cassette
15 d'expression comportant une unité d'ADN codant soit pour un précurseur d'un variant gp160 hybride non-clivable et soluble, soit pour un variant gp160 hybride non-clivable et soluble ; ledit variant ayant une séquence en acides aminés comprenant :

- 20 i) une première région dérivée de la gp160 d'une première souche du virus HIV et située dans cette dernière gp160 de l'acide aminé en position X à l'acide aminé en position Y, X étant un nombre de 1 à 271 et Y étant un nombre de 306 à 482 ;
- 25 ii) une deuxième région dérivée d'un variant gp160 soluble, non-clivable de type A ayant pour origine une deuxième souche du virus HIV et située dans ce dernier variant gp160 de l'acide aminé en position Y + 1 à l'acide aminé en position C-terminale, Y étant tel que défini ci-dessus ; et
- 30 iii) quand X est différent de 1, une troisième région dérivée de la gp160 de ladite deuxième souche du virus HIV et située dans cette dernière gp160 de l'acide aminé en position 1 à l'acide aminé en position X - 1, X étant tel que défini ci-dessus ;

ladite méthode comprenant :

35

- a) l'acte de cloner un fragment d'ADN codant pour ladite première région ; et
- b) l'acte d'insérer le fragment d'ADN cloné en a) dans un site d'une cassette qui comprend :

40

- 7 -

m) en amont du site et en séquence :

- i) un promoteur,
- 5 ii) un codon d'initiation de traduction,
- iii) en option, une première région d'ADN codant pour un peptide signal, et
- 10 iv) quand X est tel que défini ci-dessus mais différent de 1, une deuxième région d'ADN codant pour ladite troisième région de la séquence en acides aminés dudit variant gp160 hybride ; et

n) en aval du site et en séquence :

- 15 i) quand Y est tel que défini ci-dessus, une troisième région d'ADN codant pour ladite deuxième région de la séquence en acides aminés dudit variant gp160 hybride et
- 20 ii) un codon de terminaison de traduction ;

pour obtenir une cassette d'expression comportant ladite unité d'ADN.

Une telle cassette d'expression est l'outil indispensable qui permet de produire un
25 variant gp160 hybride, soluble et non-clivable dans un système d'expression hétérologue.

Dans la méthode selon l'invention, le fragment d'ADN codant pour ladite première région peut être cloné selon n'importe quelle méthode en usage. On indique cependant que ce clonage peut être avantageusement effectué par la technique PCR à partir de l'ADN
30 génomique de cellules infectées par ladite deuxième souche du virus HIV. L'insertion du fragment d'ADN cloné s'effectue dans des sites de restriction initialement présents dans la cassette ou créés à cet effet, par exemple en utilisant un polylinker de manière appropriée. Une illustration de cette méthode est présentée par la suite dans les exemples.

35 Par "unité d'ADN codant pour un polypeptide quelconque", on entend un segment d'ADN dont le premier codon en position 5' et le dernier codon en position 3' codent respectivement pour le premier acide aminé en position N-terminale et le dernier acide aminé en position C-terminale du polypeptide.

40 Selon un mode préféré, la méthode selon l'invention est mise en oeuvre pour

- 8 -

construire une cassette d'expression comportant une unité d'ADN codant pour un précurseur d'un variant gp160 hybride non-clivable et soluble.

Par "précurseur d'un variant gp160 selon l'invention", on signifie un polypeptide
5 comportant un peptide signal et un variant gp160 mature ; l'extrémité N-terminale dudit variant étant associée à l'extrémité C-terminale du peptide signal par liaison peptidique. Ce précurseur est le produit initial de l'expression de l'unité d'ADN. Il est en particulier présent dans le cytoplasme de la cellule hôte, associé en position N-terminale avec un résidu methionine d'initiation de traduction. Le peptide signal permet d'initier le transfert
10 du variant gp160 dans le réticulum endoplasmique. Lors de ce transfert, le peptide signal est abandonné par clivage protéolytique pour donner la forme mature du variant gp160. Dans la suite du texte, le terme "variant gp160" fait exclusivement référence à sa forme mature.

15 Le peptide signal peut être n'importe quel peptide signal en usage. Il doit être cependant choisi en tenant compte de l'organisme-hôte destiné à la production du variant gp160 selon l'invention. Par exemple, si l'organisme-hôte est une cellule de mammifère, le peptide signal peut être avantageusement sélectionné parmi les peptides signal des précurseurs des gp160 des différentes souches du virus HIV, indépendamment de l'origine
20 des première, deuxième et troisième régions du variant selon l'invention. De manière alternative, on peut utiliser des peptides signal synthétiques. De tels peptides signal sont par exemple, des peptides signal hybrides dont l'extrémité N-terminale dérive du précurseur de la gp160 de ladite deuxième souche du virus HIV et dont l'extrémité C-terminale dérive du précurseur de la gp160 de ladite première souche du virus HIV. A
25 titre indicatif, on mentionne de plus, que le peptide signal du précurseur de la glycoprotéine d'une souche du virus de la rage peut être aussi utilisé. Enfin, l'homme du métier doit comprendre que cette liste est non-limitative.

L'invention propose de même un variant gp160 hybride, non-clivable et soluble, qui
30 comprend :

- i) une première région dérivée de la gp160 d'une première souche du virus HIV et située dans cette dernière gp160 de l'acide aminé en position X à l'acide aminé en position Y, X étant un nombre de 1 à 271 et Y étant un nombre de
35 306 à 482 ;
- ii) une deuxième région dérivée d'un variant gp160 soluble, non-clivable de type A ayant pour origine une deuxième souche du virus HIV et située dans ce dernier variant gp160 de l'acide aminé en position Y + 1 à l'acide aminé en
40 position C-terminale, Y étant tel que défini ci-dessus ; et

- iii) quand X est différent de 1, une troisième région dérivée de la gp160 de ladite deuxième souche du virus HIV et située dans cette dernière gp160 de l'acide aminé en position 1 à l'acide aminé en position X - 1, X étant tel que défini ci-dessus.

5

La première, ainsi que la deuxième et troisième régions du variant gp160 selon l'invention peuvent être dérivées de la gp160 de n'importe quelle souche du virus HIV, à condition que la première souche soit différente de la deuxième. De manière préférée, la première région dérive de la gp160 d'une souche virale sélectionnée parmi les souches
10 HIV-2 Rod, HIV-1 Eli, HIV-1 RF, HIV-1 SF2C, HIV-1 SC et HIV-1 MN, cette dernière étant plus particulièrement préférée. De même, la deuxième et troisième régions ont de préférence pour origine la gp160 de la souche HIV-1 Bru.

Par "variant gp160 non-clivable de type A", on entend un variant gp160 :

- 15 - soit dérivé d'une gp160 d'une souche virale de type HIV-1 et ne contenant plus le site de clivage majeur,
- soit dérivé d'une gp160 d'une souche virale de type HIV-2 et ne contenant plus le site de clivage.

Lorsque le variant gp160 dérive d'une gp160 d'une souche virale de type HIV-1, ce
20 variant est de préférence de type A' ; c'est-à-dire ne contenant ni le site de clivage majeur, ni le site de clivage mineur.

Par "variant gp160 soluble", on entend un variant gp160 dérivé d'une gp160 native qui ne contient plus de région transmembranaire ou dont la région transmembranaire
25 native a été mutée de manière à ne plus pouvoir assurer sa fonction d'ancrage dans la membrane. De plus, un tel variant soluble peut avantageusement ne plus contenir de région hydrophobe.

Par "cassette d'expression", on entend un segment d'ADN comprenant une unité
30 d'ADN à exprimer ainsi que les éléments nécessaires à l'expression de cette dernière. L'expression d'une unité d'ADN est obtenue par transcription de cette unité d'ADN en ARN messager et par traduction de cet ARN messager en protéine. Par conséquent, les éléments indispensables sont un promoteur constitutif ou inductible (initiation de la transcription), un codon d'initiation de traduction (codon ATG) et un codon de fin de
35 traduction (codon TAG, TAA ou TGA). En outre, une cassette d'expression peut aussi contenir d'autres éléments ; par exemple, un terminateur de transcription.

L'homme du métier sait bien évidemment choisir le promoteur et le terminateur appropriés en fonction de l'organisme-hôte dans lequel on souhaite exprimer l'unité
40 d'ADN et en fonction du vecteur dans lequel la cassette d'expression doit être insérée

- 10 -

pour assurer sa réplication.

En accord avec ce qui précède, l'invention propose de même, une cassette d'expression comportant une unité d'ADN codant soit pour un précurseur d'un variant
5 gp160 selon l'invention, soit pour un variant gp160 selon l'invention ainsi que les éléments nécessaires à expression de ladite unité d'ADN.

Afin d'assurer sa réplication autonome dans un organisme hôte, une cassette d'expression selon l'invention peut être insérée dans différents types de vecteur ayant une
10 origine de réplication adaptée à l'organisme hôte ; par exemple un plasmide ou un virus. Les vecteurs de type viral ont en particulier la capacité d'intégrer dans leur génome une quantité substantielle d'ADN étranger sans nuire à leur capacité de réplication. Parmi ceux-ci, on compte par exemple, les poxvirus, tels que le virus de la vaccine, le poxvirus du canari et la poxvirus de la variole aviaire ; les baculovirus tels que et les adénovirus
15 tel que l'Adénovirus-2 ou l'Adénovirus-5.

Outre leur utilisation dans un système de production hétérologue, certains de ces vecteurs de type viral peuvent être fonctionnels à titre d'agent de vaccination. Il s'agit en
20 particulier des poxvirus et des adénovirus.

C'est pourquoi, l'invention concerne de même un vecteur viral dans le génome duquel est inséré une cassette d'expression selon l'invention.

Sous un autre aspect de l'invention, on procure aussi une cellule transformée par une
25 cassette d'expression selon l'invention. La cassette d'expression qui transforme la cellule peut être véhiculée par un plasmide ou, de manière alternative, être intégrée dans le génome de la cellule hôte.

L'organisme-hôte destiné à la production du variant gp160 selon l'invention peut
30 être n'importe quel type de cellule, de préférence eucaryote. Ceci inclut par exemple : des champignons tels que les levures, des cellules d'insectes et des cellules de mammifères.

L'invention propose, en outre, deux procédés alternatifs visant à la production d'un variant gp160 selon l'invention :

35 - le premier procédé comprend l'acte de cultiver une cellule selon l'invention et l'acte de récolter ledit variant à partir de la culture,

- le second procédé comprend l'acte d'infecter une culture de cellules avec un
40 vecteur viral selon l'invention et l'acte de récolter ledit variant à partir de la culture.

- 11 -

Un variant gp160 ainsi qu'un vecteur viral selon l'invention possèdent une activité vaccinale à l'encontre d'un virus HIV et, par conséquent, sont utiles à titre de produits pharmaceutiques en particulier destinés au traitement ou à la prévention du Sida.

5 En conséquence, l'invention propose :

i) une composition pharmaceutique destinée au traitement curatif ou préventif du Sida qui comprend à titre d'agent thérapeutique au moins un variant gp160 ou un vecteur viral selon l'invention,

10

ii) une méthode de traitement curatif ou préventif du Sida qui comprend l'acte d'administrer une quantité thérapeutiquement effective d'un variant gp160 ou d'un vecteur viral selon l'invention à un patient ayant besoin d'un tel traitement.

15 iii) l'usage d'un variant gp160 ou d'un vecteur viral selon l'invention à titre d'agent thérapeutique destiné au traitement curatif ou préventif du Sida.

De manière préférée, une composition pharmaceutique selon l'invention peut contenir plusieurs variants gp160 selon l'invention ; chaque variant gp160 possédant au moins un troisième domaine hypervariable (boucle V_3) différent des autres variants présents dans la composition. Une telle composition pharmaceutique devrait donc permettre de protéger correctement un individu vis-à-vis de diverses souches HIV.

25 Une composition pharmaceutique selon l'invention peut contenir en outre d'autres agents thérapeutiques comme, par exemple, un peptide correspondant essentiellement au troisième domaine hypervariable d'une gp160 (ci-après appelé peptide V_3). De préférence, un tel peptide a une séquence substantiellement identique à la séquence du troisième domaine d'un variant gp160 contenu dans la composition pharmaceutique selon l'invention. De manière similaire, plusieurs peptides V_3 peuvent être présents dans la composition. En particulier, si la composition pharmaceutique selon l'invention contient 30 différents variants gp160, on peut bien évidemment y ajouter les peptides V_3 correspondants.

Une composition pharmaceutique selon l'invention peut être fabriquée de manière conventionnelle. En particulier, on associe un variant gp160 selon l'invention avec un diluant ou un support acceptable d'un point de vue pharmaceutique. Enfin, une composition selon l'invention peut contenir un adjuvant de vaccination tel que l'alun. De manière alternative, cet adjuvant peut être ajouté à une composition selon l'invention juste avant usage.

40

- 12 -

Une composition selon l'invention peut être administrée par n'importe quelle voie conventionnelle en usage dans le domaine des vaccins, en particulier par voie sous-cutanée, par exemple sous forme de solution ou de suspension injectable. L'administration peut avoir lieu en dose unique ou répétée une ou plusieurs fois après un certain délai d'intervalle. Le dosage approprié varie en fonction de divers paramètres, par exemple, de l'individu traité ou du mode d'administration.

De manière alternative, une composition pharmaceutique selon l'invention peut être présentée comme partie d'un kit de traitement. A titre d'exemple, on indique qu'un tel kit peut contenir :

- d'une part, une composition pharmaceutique contenant au moins un variant gp160 selon l'invention et,

- d'autre part, une composition pharmaceutique contenant au moins un peptide V₃,

- ainsi qu'une notice spécifiant les instructions relatives à l'administration séquentielle ou concomitante des compositions pharmaceutiques contenues dans le kit.

20

L'invention est illustrée ci-après par référence aux figures suivantes:

La Figure 1 présente la séquence des acides aminés des précurseurs des gp160 natives des souches virales HIV-1 Bru (a), HIV-1 MN (b), HIV-1 Eli (c), HIV-1 RF (d), HIV-1 SC (e) et HIV-1 SF2C (f). Les résidus acides aminés des séquences signal sont numérotés de la position -1 à -29. Le résidu méthionine en position -30 correspond au codon d'initiation de traduction. Les résidus acides aminés des gp160 matures sont numérotés de la position 1 à 841. L'astérisque (*) symbolise l'identité des résidus acides aminés à une position donnée tandis que le point (.) indique un changement conservatif (acides aminés différents, mais appartenant à une même classe).

La Figure 2 présente la séquence des acides aminés des précurseurs des gp160 natives des souches virales HIV-1 Bru (a) et HIV-2 Rod (b). Les résidus acides aminés des séquences signal sont numérotés de la position -1 à -29. Le résidu méthionine en position -30 correspond au codon d'initiation de traduction. Les résidus acides aminés des gp160 matures sont numérotés de la position 1 à 841. L'astérisque (*) symbolise l'identité des résidus acides aminés à une position donnée tandis que le point (.) indique un changement conservatif (acides aminés différents, mais appartenant à une même classe).

La Figure 3 présente la séquence nucléotidique du fragment *PstI-PstI* du vecteur

- 13 -

pTG1163 qui comporte la séquence codant pour un précurseur d'un variant gp160-Bru non-clivable soluble (domaine transmembranaire absent) ainsi que la séquence en acides aminés de ce précurseur.

5 La Figure 4 schématise les étapes de la construction des bactériophages M13TG4168 et M13TG4174.

La Figure 5 présente la séquence nucléotidique codant pour le précurseur de la gp120-MN ainsi que la séquence en acides aminés de ce précurseur. Les oligonucléotides
10 OTG2624 et OTG2625 destinés à l'amplification d'un fragment d'ADN codant au moins pour le troisième domaine hypervariable de la gp120-MN sont représentés au dessus de leur région d'hybridation.

La Figure 6 présente la séquence nucléotidique codant pour le précurseur de la
15 gp120-Eli ainsi que la séquence en acides aminés de ce précurseur. Les oligonucléotides OTG2624 et OTG2625 destinés à l'amplification d'un fragment d'ADN codant au moins pour le troisième domaine hypervariable de la gp120-MN sont représentés au dessus de leur région d'hybridation.

20 La Figure 7 présente la séquence nucléotidique codant pour le précurseur de la gp120-RF ainsi que la séquence en acides aminés de ce précurseur. Les oligonucléotides OTG2624 et OTG2625 destinés à l'amplification d'un fragment d'ADN codant au moins pour le troisième domaine hypervariable de la gp120-RF sont représentés au dessus de leur région d'hybridation.

25 La Figure 8 présente la séquence nucléotidique codant pour le précurseur de la gp120-SF2C ainsi que la séquence en acides aminés de ce précurseur. Les oligonucléotides OTG2624 et OTG2625 destinés à l'amplification d'un fragment d'ADN codant au moins
30 leur région d'hybridation.

Dans les exemples suivants, pour faciliter l'écriture et la compréhension, on entend par "séquence signal", une séquence signal incluant le résidu méthionine d'initiation de traduction.

35

Exemple 1 : Construction d'une cassette d'insertion porté par le bactériophage M13TG4168.

5 Tel que montré à la Figure 3, le fragment d'ADN *PstI-PstI* du plasmide pTG1163 décrit dans la demande de brevet EPA 245 136 comporte une séquence d'ADN codant pour un précurseur d'un variant gp160-Bru non-clivable et soluble. Ce fragment d'ADN *PstI-PstI* est inséré dans le bactériophage M13mp701 (décrit par M.P. Kieny et al. Gene (1983) 26 : 91) préalablement digéré par *PstI* pour donner le bactériophage M13TG4137
10 (Figure 4). La numérotation des nucléotides du fragment *PstI-PstI* telle que indiquée à la Figure 3 sert de référence dans la suite de l'exemple 1.

Le plasmide pTG1163 est coupé par *PstI* et *KpnI* et le fragment d'ADN correspondant aux nucléotides 1 à 138 (Figure 3) est inséré dans le bactériophage
15 M13TG130 (décrit par M.P. Kieny et al. (1983), supra) préalablement digéré par *PstI* et *KpnI*. On obtient ainsi le bactériophage M13TG4147 (Figure 4).

Le bactériophage M13TG4137 est coupé par *BglII*, traité par l'enzyme klenow (Boehringer Mannheim) pour obtenir des extrémités franches, puis coupé par *EcoRI*. Le
20 fragment *EcoRI-BglII*° issu de cette digestion et comportant la séquence correspondant aux nucléotides 1424 à 2644 (Figure 3) est inséré dans le bactériophage M13TG4147 préalablement digéré par *EcoRV* et *EcoRI*. On obtient ainsi le bactériophage M13TG4158 qui comporte :

- 25 i) un fragment d'ADN correspondant aux nucléotides 1 à 138,
- ii) la séquence restante du polylinker du bactériophage M13TG4147 c'est à dire ATCGCATGCG,
- 30 iii) un fragment d'ADN correspondant aux nucléotides 1424 à 2644.

Le bactériophage M13TG4158 simple brin anti-sens, sert de matrice pour une mutation-délétion effectuée en utilisant la trousse Amersham et l'oligonucléotide OTG2623 dont la séquence est la suivante :

35

┌───*SphI*───┐ ┌───*SmaI*───┐

GGGGGTGGAAATGGGGCA GCATGC AT CCCGGG CACAGAATCACGTGGTGC.

40 Cette mutagenèse permet simultanément de déléter un fragment de 184 paires de

- 15 -

bases qui comporte les nucléotides 67 à 138, la séquence ATCGCATGCG et les nucléotides 1424 à 1525 et de créer les sites de coupure pour les enzymes *SphI* en 5' et *SmaI* en 3' qui sont insérés entre les nucléotides 66 et 1526. On obtient ainsi le bactériophage M13TG4168 qui comprend :

5

i) un fragment d'ADN correspondant aux nucléotides 1 à 66,

ii) la séquence d'ADN GCATGCATCCCG comportant les sites de clivage des enzymes *SphI* et *SmaI*, et

10

iii) un fragment d'ADN correspondant aux nucléotides 1526 à 2644.

15

Exemple 2 : Construction d'une cassette d'expression destinée à la synthèse d'un variant gp160 hybride Bru-MN.

Tel que montré à la Figure 5, le fragment d'ADN codant pour la majeure partie d'un précurseur de la gp120-MN est cloné par la technique d'amplification génique PCR (Polymérase Chain Reaction) mise en oeuvre à partir de l'ADN génomique de cellules
20 humaines CEM infectées par la souche virale HIV1-MN. Ce clonage est effectué à l'aide des oligonucléotides OTG2624 et OTG2625 qui introduisent respectivement un site de coupure pour l'enzyme *SphI* en 5' du fragment d'ADN amplifié et un site de coupure pour l'enzyme *SmaI* en 3' du fragment d'ADN amplifié. Les séquences de ces oligonucléotides sont les suivantes :

25

┌───*SphI*───┐

OTG2624 : GGCA GCATGC TCCTTGGGATATTGATGATCTG

30

┌───*SmaI*───┐

OTG2625 : CTTTG CCCGGG TGGGTGCTACTCCTAATGGTTC

Le fragment d'ADN *SphI-SmaI* amplifié correspond aux nucléotides 53 à 1505
35 représentés à la Figure 5, compte tenu des modifications de séquence apportées par les oligonucléotides OTG2624 et OTG2625. Ce fragment est inséré dans le bactériophage M13TG4168 coupé par *SphI* et *SmaI*. On obtient ainsi le bactériophage M13TG4174 qui comporte un fragment d'ADN codant pour un précurseur d'une protéine gp160 hybride Bru-MN. Le fragment *PstI* du bactériophage M13TG4174 qui code pour le variant gp160
40 hybride Bru-MN est inséré dans le plasmide de transfert pTG186poly (décrit par M.P. Kieny et al., Biotechnology, (1986) 4 : 790), en aval du promoteur E7.5k et à l'intérieur

- 16 -

du gène du virus de la vaccine codant pour la thymidine kinase. On obtient ainsi le plasmide pTG5156.

Le plasmide pTG5156 est utilisé par la suite pour transférer le bloc d'expression de la protéine gp160 hybride Bru-MN dans le génome du virus de la vaccine, souche Copenhague, selon la méthode décrite par M.P. Kieny et al. Nature (1984), 312, 163-166). On obtient ainsi le vecteur de la vaccine VVTG5156.

10 Exemple 3 : Construction d'une cassette d'expression destinée à la synthèse d'un variant gp160 hybride Bru-Eli.

Le plasmide pTG186poly est coupé par *BamHI*, traité par l'enzyme Klenow, coupé par *SmaI* pour déléter la majeure partie du polylinker, puis reliqué sur lui-même pour donner le plasmide pTG186PE. Le polylinker ne conserve plus que les sites de clivage pour les enzymes *PstI*, *Sall* et *EcoRI*.

Le fragment d'ADN *PstI-PstI* du bactériophage M13TG4168 décrit, ci-dessus, qui comporte :

- 20 i) un fragment d'ADN correspondant aux nucléotides 1 à 66 du fragment d'ADN représenté sur la Figure 3,
- 25 ii) la séquence d'ADN GCATGCATCCCG comportant les sites de clivage des enzymes *SphI* et *SmaI*, et
- iii) un fragment d'ADN correspondant aux nucléotides 1526 à 2644 du fragment d'ADN représenté à la Figure 3,

30 est inséré dans le plasmide pTG186PE préalablement coupé par *PstI*, en aval du promoteur E7.5k et à l'intérieur du gène du virus de la vaccine codant pour la thymidine kinase. On obtient ainsi le plasmide pTG5160.

Tel que montré à la Figure 6, le fragment d'ADN codant pour la majeure partie d'un précurseur de la gp120-Eli est cloné par la technique d'amplification génique PCR mise en oeuvre à partir de l'ADN génomique de cellules humaines CEM infectées par du virus HIV1-Eli. Ce clonage est effectué à l'aide des oligonucléotides OTG2624 et OTG2625 qui introduisent respectivement un site coupure pour l'enzyme *SphI* en 5' du fragment d'ADN amplifié et un site de coupure pour l'enzyme *SmaI* en 3' du fragment d'ADN amplifié.

40 Le fragment d'ADN *SphI-SmaI* amplifié correspond aux nucléotides 53 à 1490 représentés

- 17 -

à la Figure 6, compte tenu des modifications de séquence apportées par les oligonucléotides OTG2624 et OTG2625. Ce fragment est inséré dans le bactériophage M13TG131 (décrit par M.P. Kieny et al. (1983), supra) préalablement coupé par *SphI* et *SmaI* pour donner le bactériophage M13TG4197. Puis le fragment *SphI-SmaI* du bactériophage M13TG4197 est inséré dans le plasmide de transfert pTG5160 (décrit ci-dessus) préalablement coupé par *SphI* et *SmaI*. On obtient ainsi le pTG5193.

Le plasmide pTG5193 est utilisé par la suite pour transférer le bloc d'expression de la protéine gp160 hybride Bru-Eli dans le génome du virus de la vaccine, souche Copenhague, selon la méthode décrite par M.P. Kieny et al. (1984) supra. On obtient ainsi le vecteur de la vaccine VVTG5193.

Exemple 4 : Construction d'une cassette d'expression destinée à la synthèse d'un variant gp160 hybride Bru-RF.

Tel que montré à la Figure 7, le fragment d'ADN codant pour la majeure partie d'un précurseur d'une protéine gp120-RF est cloné par la technique d'amplification génique PCR mise en oeuvre à partir de l'ADN génomique de cellules humaines CEM infectées par du virus HIV1-RF. Ce clonage est effectué à l'aide des oligonucléotides OTG2624 et OTG2625 qui introduisent respectivement un site de coupure pour l'enzyme *SphI* en 5' du fragment d'ADN amplifié et un site de coupure pour l'enzyme *SmaI* en 3' du fragment d'ADN amplifié. Le fragment d'ADN *SphI-SmaI* amplifié correspond aux nucléotides 53 à 1523 représentés à la Figure 7, compte tenu des modifications de séquence apportées par les oligonucléotides OTG2624 et OTG2625. Ce fragment est inséré dans le bactériophage M13TG131 préalablement coupé par *SphI* et *SmaI* pour donner le bactériophage M13TG4198. Puis le fragment *SphI-SmaI* du bactériophage M13TG4198 est inséré dans le plasmide de transfert pTG5160 (décrit à l'exemple 3) préalablement coupé par *SphI* et *SmaI*. On obtient ainsi le pTG5194.

Le plasmide pTG5194 est utilisé par la suite pour transférer le bloc d'expression de la protéine gp160 hybride Bru-RF dans le génome du virus de la vaccine, souche Copenhague, selon la méthode décrite par M.P. Kieny et al. (1984) supra. On obtient ainsi le vecteur de la vaccine VVTG5194.

Exemple 5 : Construction d'une cassette d'expression destinée à la synthèse d'un variant gp160 hybride Bru-SF2C.

Tel que montré à la Figure 8, le fragment d'ADN codant pour la majeure partie d'un

- 18 -

précurseur de la gp120-SF2C est cloné par la technique d'amplification génique PCR mise en oeuvre à partir de l'ADN génomique de cellules humaines CEM infectées par du virus HIV1-SF2C. Ce clonage est effectué à l'aide des oligonucléotides OTG2624 et OTG2625 qui introduisent respectivement un site de coupure pour l'enzyme *SphI* en 5' du fragment d'ADN amplifié et un site de coupure pour l'enzyme *SmaI* en 3' du fragment d'ADN amplifié. Le fragment d'ADN *SphI-SmaI* amplifié correspond aux nucléotides 53 à 1493 représentés à la Figure 8, compte tenu des modifications de séquence apportées par les oligonucléotides OTG2624 et OTG2625. Puis ce fragment est inséré dans le bactériophage M13TG131 préalablement coupé par *SphI* et *SmaI* pour donner le bactériophage M13TG4199. Le fragment *SphI-SmaI* du bactériophage M13TG4199 est inséré dans le plasmide de transfert pTG5160 (décrit à l'exemple 3) préalablement coupé par *SphI* et *SmaI*. On obtient ainsi le pTG5195.

Le plasmide pTG5195 est utilisé par la suite pour transférer le bloc d'expression de la protéine gp160 hybride Bru-SF2C dans le génome du virus de la vaccine, souche Copenhagen, selon la méthode décrite par M.P. Kieny et al. (1984) supra. On obtient ainsi le vecteur de la vaccine VVTG5195.

Exemples 6 à 9 : Production et purification des gp160 hybrides Bru-MN, Bru-Eli, Bru-RF et Bru SF2C :

Des cellules BHK-21 sont cultivées dans un milieu GMEM (Gibco) supplémenté avec 10% de sérum de veau foetal (SVF). Lorsque les cellules sont à confluence (5,8x10⁶cellules/ml), le milieu de culture est enlevé et le tapis cellulaire est lavé 2 fois avec 50 ml de PBS (Dabelcco's phosphate buffer salt ; Seromed). Puis du milieu GMEM frais sans SVF est rajouté. On ajoute alors un des virus de la vaccine VVTG5156, VVTG5193, VVTG5194 et VVTG5195 décrits ci-dessus dans les exemples 2 à 5, à une infectivité de 1 ufp/cellule (unité formant plages), et l'infection est poursuivie pendant 72 h. Le lysat est alors centrifugé 20 min. à 10 000 g pour éliminer les débris cellulaires et on récupère le surnageant contenant entre autre une gp160 hybride et des virions.

A 20 ml de surnageant de culture obtenu comme précédemment décrit, on ajoute 1,3 ml d'une solution de chlorure de zinc (ZnCl₂) à 1 M pour avoir une concentration finale en ZnCl₂ de 60 mM. Le mélange est laissé 1 h dans de la glace ; puis le surnageant de précipitation est récupéré après centrifugation à 3000 t/min. pendant 20 min. dans une centrifugeuse Minifuge RF Heraeus. Cette méthode élimine par précipitation les virions ainsi que la majorité des protéines contaminantes. Dans chaque cas, on obtient ainsi une solution purifiée du variant gp160 hybride.

REVENDICATIONS

1. Une méthode pour construire une cassette d'expression comportant une unité d'ADN codant soit pour un précurseur d'un variant gp160 hybride, non-clivable et soluble, soit pour un variant gp160 hybride, non-clivable et soluble ; ledit variant gp160 ayant une séquence en acides aminés comprenant :
- 5
- i) une première région dérivée de la gp160 d'une première souche du virus HIV et située dans cette dernière gp160 de l'acide aminé en position X à l'acide aminé en position Y, X étant un nombre de 1 à 271 et Y étant un nombre de 306 à 482 ;
- 10
- ii) une deuxième région dérivée d'un variant gp160 soluble, non-clivable de type A ayant pour origine une deuxième souche du virus HIV et située dans ce dernier variant gp160 de l'acide aminé en position Y + 1 à l'acide aminé en position C-terminale, Y étant tel que défini ci-dessus ; et
- 15
- iii) quand X est différent de 1, une troisième région dérivée de la gp160 de ladite deuxième souche du virus HIV et située dans cette dernière gp160 de l'acide aminé en position 1 à l'acide aminé en position X - 1, X étant tel que défini ci-dessus ;
- 20
- ladite méthode comprenant :
- a) l'acte de cloner un fragment d'ADN codant pour ladite première région ; et
- 25
- b) l'acte d'insérer le fragment d'ADN cloné en a) dans un site d'une cassette qui comprend :
- m) en amont du site et en séquence :
- 30
- i) un promoteur,
- ii) un codon d'initiation de traduction,
- iii) en option, une première région d'ADN codant pour un peptide signal, et
- 35
- iv) quand X est tel que défini ci-dessus mais différent de 1. une

- 20 -

deuxième région d'ADN codant pour ladite troisième région de la séquence en acides aminés dudit variant gp160 hybride ; et

n) en aval du site et en séquence :

i) quand Y est tel que défini ci-dessus, une troisième région d'ADN codant pour ladite deuxième région de la séquence en acides aminés dudit variant gp160 hybride et

ii) un codon de terminaison de traduction ;

pour obtenir une cassette d'expression comportant ladite unité d'ADN.

2. Une méthode selon la revendication 1, pour construire une cassette d'expression comportant une unité d'ADN codant pour un précurseur d'un variant gp160 hybride, non-clivable et soluble, ayant une séquence en acides aminés qui comprend :

i) une première région dérivée de la gp160 d'une première souche du virus HIV et située dans cette dernière gp160 de l'acide aminé en position X à l'acide aminé en position Y, X étant un nombre de 1 à 271 et Y étant un nombre de 306 à 482 ;

ii) une deuxième région dérivée d'un variant gp160 soluble, non-clivable de type A ayant pour origine une deuxième souche du virus HIV et située dans ce dernier variant gp160 de l'acide aminé en position Y + 1 à l'acide aminé en position C-terminale, Y étant tel que défini ci-dessus ; et

iii) quand X est différent de 1, une troisième région dérivée de la gp160 de ladite deuxième souche du virus HIV et située dans cette dernière gp160 de l'acide aminé en position 1 à l'acide aminé en position X - 1, X étant tel que défini ci-dessus ;

ladite méthode comprenant :

a) l'acte de cloner un fragment d'ADN codant pour ladite première région ; et

b) l'acte d'insérer le fragment d'ADN cloné en a) dans un site d'une cassette qui

- 21 -

m) en amont du site et en séquence :

- i) un promoteur,
- ii) un codon d'initiation de traduction,
- 5 iii) une première région d'ADN codant pour un peptide signal, et
- iv) quand X est tel que défini ci-dessus mais différent de 1, une deuxième région d'ADN codant pour ladite troisième région de la séquence en acides aminés dudit variant gp160 hybride ; et

10 n) en aval du site et en séquence :

- i) quand Y est tel que défini ci-dessus, une troisième région d'ADN codant pour ladite deuxième région de la séquence en acides aminés dudit variant gp160 hybride et
- 15 ii) un codon de terminaison de traduction ;

pour obtenir une cassette d'expression comportant ladite unité d'ADN.

20 3. Une méthode selon la revendication 1, pour construire une cassette d'expression comportant une unité d'ADN codant soit pour un précurseur d'un variant gp160 hybride, non-clivable et soluble, soit pour un variant gp160 hybride, non-clivable et soluble ; ledit variant gp160 ayant une séquence en acides aminés comprenant :

- 25 i) une première région dérivée de la gp160 d'une première souche du virus HIV et située dans cette dernière gp160 de l'acide aminé en position X à l'acide aminé en position Y, X étant un nombre de 1 à 271 et Y étant un nombre de 306 à 476 ;
- 30 ii) une deuxième région dérivée d'un variant gp160 soluble, non-clivable de type A' ayant pour origine une deuxième souche du virus HIV et située dans ce dernier variant gp160 de l'acide aminé en position Y + 1 à l'acide aminé en position C-terminale, Y étant tel que défini ci-dessus ; et
- 35 iii) quand X est différent de 1, une troisième région dérivée de la gp160 de ladite deuxième souche du virus HIV et située dans cette dernière gp160 de l'acide

- 22 -

aminé en position 1 à l'acide aminé en position X - 1, X étant tel que défini ci-dessus ;

ladite méthode comprenant :

- 5 a) l'acte de cloner un fragment d'ADN codant pour ladite première région ; et
- b) l'acte d'insérer le fragment d'ADN cloné en a) dans un site d'une cassette qui comprend :
- 10 m) en amont du site et en séquence :
- i) un promoteur,
- 15 ii) un codon d'initiation de traduction,
- iii) en option, une première région d'ADN codant pour un peptide signal, et
- 20 iv) quand X est tel que défini ci-dessus mais différent de 1, une deuxième région d'ADN codant pour ladite troisième région de la séquence en acides aminés dudit variant gp160 hybride ; et
- n) en aval du site et en séquence :
- 25 i) quand Y est tel que défini ci-dessus, une troisième région d'ADN codant pour ladite deuxième région de la séquence en acides aminés dudit variant gp160 hybride et
- ii) un codon de terminaison de traduction ;

pour obtenir une cassette d'expression comportant ladite unité d'ADN.

30

4. Une méthode selon la revendication 1 ou 3, pour construire une cassette d'expression comportant une unité d'ADN codant soit pour un précurseur d'un variant gp160 hybride, non-clivable et soluble, soit pour un variant gp160 hybride, non-clivable et soluble ; ledit variant ayant une séquence en acides aminés
- 35 comprenant :

- 23 -

- i) une première région dérivée de la gp160 d'une première souche du virus HIV et située dans cette dernière gp160 de l'acide aminé en position X à l'acide aminé en position Y, X étant un nombre de 1 à 271 et Y étant un nombre de 306 à 482 ;
- 5 ii) une deuxième région dérivée d'un variant gp160 soluble, non-clivable de type A ayant pour origine une deuxième souche du virus HIV et située dans ce dernier variant gp160 de l'acide aminé en position Y + 1 à l'acide aminé en position C-terminale, Y étant tel que défini ci-dessus et ladite deuxième souche étant la souche HIV1-Bru ; et
- 10 iii) quand X est différent de 1, une troisième région dérivée de la gp160 de ladite deuxième souche du virus HIV et située dans cette dernière gp160 de l'acide aminé en position 1 à l'acide aminé en position X - 1, X étant tel que défini ci-dessus ;
- 15 ladite méthode comprenant :
- a) l'acte de cloner un fragment d'ADN codant pour ladite première région ; et
- b) l'acte d'insérer le fragment d'ADN cloné en a) dans un site d'une cassette qui
- 20 comprend :
- m) en amont du site et en séquence :
- i) un promoteur,
- 25 ii) un codon d'initiation de traduction,
- iii) en option, une première région d'ADN codant pour un peptide signal, et
- 30 iv) quand X est tel que défini ci-dessus mais différent de 1, une deuxième région d'ADN codant pour ladite troisième région de la séquence en acides aminés dudit variant gp160 hybride ; et
- n) en aval du site et en séquence :
- 35 i) quand Y est tel que défini ci-dessus, une troisième région d'ADN codant pour ladite deuxième région de la séquence en acides

- 24 -

aminés dudit variant gp160 hybride et

ii) un codon de terminaison de traduction ;

pour obtenir une cassette d'expression comportant ladite unité d'ADN.

5

5. Une méthode selon n'importe laquelle des revendications 1, 3 et 4, pour construire une cassette d'expression comportant une unité d'ADN codant soit pour un précurseur d'un variant gp160 hybride, non-clivable et soluble, soit pour un variant gp160 hybride, non-clivable et soluble ; ledit variant ayant une séquence en acides aminés comprenant :

10

15

i) une première région dérivée de la gp160 d'une première souche du virus HIV et située dans cette dernière gp160 de l'acide aminé en position X à l'acide aminé en position Y, X étant un nombre de 1 à 271 et Y étant un nombre de 306 à 482 et ladite première souche étant sélectionnée parmi les souches HIV-1 MN, HIV-1 Eli, HIV-1 RF, HIV-1 SF2C et HIV-1 SC;

20

ii) une deuxième région dérivée d'un variant gp160 soluble, non-clivable de type A ayant pour origine une deuxième souche du virus HIV et située dans ce dernier variant gp160 de l'acide aminé en position Y + 1 à l'acide aminé en position C-terminale, Y étant tel que défini ci-dessus ; et

25

iii) quand X est différent de 1, une troisième région dérivée de la gp160 de ladite deuxième souche du virus HIV et située dans cette dernière gp160 de l'acide aminé en position 1 à l'acide aminé en position X - 1, X étant tel que défini ci-dessus ;

ladite méthode comprenant :

30

a) l'acte de cloner un fragment d'ADN codant pour ladite première région ; et

b) l'acte d'insérer le fragment d'ADN cloné en a) dans un site d'une cassette qui comprend :

m) en amont du site et en séquence :

35

i) un promoteur,

- 25 -

- ii) un codon d'initiation de traduction,
 - iii) en option, une première région d'ADN codant pour un peptide signal, et
 - iv) quand X est tel que défini ci-dessus mais différent de 1, une
5 deuxième région d'ADN codant pour ladite troisième région de la séquence en acides aminés dudit variant gp160 hybride ; et
- n) en aval du site et en séquence :
- i) quand Y est tel que défini ci-dessus, une troisième région d'ADN
10 codant pour ladite deuxième région de la séquence en acides aminés dudit variant gp160 hybride et
 - ii) un codon de terminaison de traduction ;
- 15 pour obtenir une cassette d'expression comportant ladite unité d'ADN.
6. Une méthode selon la revendication 5, pour construire une cassette d'expression comportant une unité d'ADN codant soit pour un précurseur d'un variant gp160
20 hybride, non-clivable et soluble, soit pour un variant gp160 hybride, non-clivable et soluble ; ledit variant ayant une séquence en acides aminés comprenant :
- i) une première région dérivée de la gp160 d'une première souche du virus HIV et située dans cette dernière gp160 de l'acide aminé en position X à l'acide
25 aminé en position Y, X étant un nombre de 1 à 271 et Y étant un nombre de 306 à 482 et ladite première souche étant la souche HIV-1 MN;
 - ii) une deuxième région dérivée d'un variant gp160 soluble, non-clivable de type A ayant pour origine une deuxième souche du virus HIV et située dans ce
30 dernier variant gp160 de l'acide aminé en position Y + 1 à l'acide aminé en position C-terminale, Y étant tel que défini ci-dessus ; et
 - iii) quand X est différent de 1, une troisième région dérivée de la gp160 de ladite
35 deuxième souche du virus HIV et située dans cette dernière gp160 de l'acide aminé en position 1 à l'acide aminé en position X - 1, X étant tel que défini ci-dessus ;

- 26 -

ladite méthode comprenant :

- a) l'acte de cloner un fragment d'ADN codant pour ladite première région ; et
- b) l'acte d'insérer le fragment d'ADN cloné en a) dans un site d'une cassette qui comprend :

5

- m) en amont du site et en séquence :

- i) un promoteur,

10

- ii) un codon d'initiation de traduction,

- iii) en option, une première région d'ADN codant pour un peptide signal, et

15

- iv) quand X est tel que défini ci-dessus mais différent de 1, une deuxième région d'ADN codant pour ladite troisième région de la séquence en acides aminés dudit variant gp160 hybride ; et

- n) en aval du site et en séquence :

20

- i) quand Y est tel que défini ci-dessus, une troisième région d'ADN codant pour ladite deuxième région de la séquence en acides aminés dudit variant gp160 hybride et

- ii) un codon de terminaison de traduction ;

25

pour obtenir une cassette d'expression comportant ladite unité d'ADN.

- 7. Une méthode selon n'importe laquelle des revendications 1 et 3 à 6, pour construire une cassette d'expression comportant une unité d'ADN codant soit pour un précurseur d'un variant gp160 hybride, non-clivable et soluble, soit pour un variant gp160 hybride, non-clivable et soluble ; ledit variant ayant une séquence en acides aminés comprenant :

30

- i) une première région dérivée de la gp160 d'une première souche du virus HIV et située dans cette dernière gp160 de l'acide aminé en position X à l'acide aminé en position Y, X étant un nombre de 1 à 271 et Y étant un nombre de

35

- 27 -

450 à 482 ;

- ii) une deuxième région dérivée d'un variant gp160 soluble, non-clivable de type A' ayant pour origine une deuxième souche du virus HIV et située dans ce dernier variant gp160 de l'acide aminé en position Y + 1 à l'acide aminé en position C-terminale, Y étant tel que défini ci-dessus ; et

- iii) quand X est différent de 1, une troisième région dérivée de la gp160 de ladite deuxième souche du virus HIV et située dans cette dernière gp160 de l'acide aminé en position 1 à l'acide aminé en position X - 1, X étant tel que défini ci-dessus ;

ladite méthode comprenant :

- a) l'acte de cloner un fragment d'ADN codant pour ladite première région ; et

- b) l'acte d'insérer le fragment d'ADN cloné en a) dans un site d'une cassette qui comprend :

m) en amont du site et en séquence :

i) un promoteur,

ii) un codon d'initiation de traduction,

iii) en option, une première région d'ADN codant pour un peptide signal, et

iv) quand X est tel que défini ci-dessus mais différent de 1, une deuxième région d'ADN codant pour ladite troisième région de la séquence en acides aminés dudit variant gp160 hybride ; et

n) en aval du site et en séquence :

i) quand Y est tel que défini ci-dessus, une troisième région d'ADN codant pour ladite deuxième région de la séquence en acides aminés dudit variant gp160 hybride et

ii) un codon de terminaison de traduction ;

- 28 -

pour obtenir une cassette d'expression comportant ladite unité d'ADN.

8. Une méthode selon n'importe laquelle des revendications 1 et 3 à 7, pour construire une cassette d'expression comportant une unité d'ADN codant soit pour un précurseur d'un variant gp160 hybride, non-clivable et soluble, soit pour un variant gp160 hybride, non-clivable et soluble ; ledit variant ayant une séquence en acides aminés comprenant :

i) une première région dérivée de la gp160 d'une première souche du virus HIV et située dans cette dernière gp160 de l'acide aminé en position X à l'acide aminé en position Y, X étant un nombre de 1 à 97 et Y étant un nombre de 306 à 482 ;

ii) une deuxième région dérivée d'un variant gp160 soluble, non-clivable de type A' ayant pour origine une deuxième souche du virus HIV et située dans ce dernier variant gp160 de l'acide aminé en position Y + 1 à l'acide aminé en position C-terminale, Y étant tel que défini ci-dessus ; et

iii) quand X est différent de 1, une troisième région dérivée de la gp160 de ladite deuxième souche du virus HIV et située dans cette dernière gp160 de l'acide aminé en position 1 à l'acide aminé en position X - 1, X étant tel que défini ci-dessus ;

ladite méthode comprenant :

a) l'acte de cloner un fragment d'ADN codant pour ladite première région ; et

b) l'acte d'insérer le fragment d'ADN cloné en a) dans un site d'une cassette qui comprend :

m) en amont du site et en séquence :

i) un promoteur,

ii) un codon d'initiation de traduction,

iii) en option, une première région d'ADN codant pour un peptide signal, et

- 29 -

iv) quand X est tel que défini ci-dessus mais différent de 1, une deuxième région d'ADN codant pour ladite troisième région de la séquence en acides aminés dudit variant gp160 hybride ; et

n) en aval du site et en séquence :

5 i) quand Y est tel que défini ci-dessus, une troisième région d'ADN codant pour ladite deuxième région de la séquence en acides aminés dudit variant gp160 hybride et

ii) un codon de terminaison de traduction ;

10

pour obtenir une cassette d'expression comportant ladite unité d'ADN.

9. Une méthode selon n'importe laquelle des revendications 1 et 3 à 8, pour construire
15 une cassette d'expression comportant une unité d'ADN codant soit pour un précurseur d'un variant gp160 hybride, non-clivable et soluble, soit pour un variant gp160 hybride, non-clivable et soluble ; ledit variant ayant une séquence en acides aminés comprenant :

20 i) une première région dérivée de la gp160 d'une première souche du virus HIV et située dans cette dernière gp160 de l'acide aminé en position X à l'acide aminé en position Y, X étant un nombre de 1 à 97 et Y étant un nombre de 306 à 482 ;

25 ii) une deuxième région dérivée d'un variant gp160 soluble, non-clivable de type A' ayant pour origine une deuxième souche du virus HIV et située dans ce dernier variant gp160 de l'acide aminé en position Y + 1 à l'acide aminé en position C-terminale, Y étant tel que défini ci-dessus ; et

30 iii) quand X est différent de 1, une troisième région dérivée de la gp160 de ladite deuxième souche du virus HIV et située dans cette dernière gp160 de l'acide aminé en position 1 à l'acide aminé en position X - 1, X étant tel que défini ci-dessus ;

ladite méthode comprenant :

35 a) l'acte de cloner un fragment d'ADN codant pour ladite première région par la technique PCR à partir de l'ADN génomique de cellules infectées par ladite

- 30 -

deuxième souche du virus HIV ; et

b) l'acte d'insérer le fragment d'ADN cloné en a) dans un site d'une cassette qui comprend :

m) en amont du site et en séquence :

i) un promoteur,

ii) un codon d'initiation de traduction,

iii) en option, une première région d'ADN codant pour un peptide signal, et

iv) quand X est tel que défini ci-dessus mais différent de 1, une deuxième région d'ADN codant pour ladite troisième région de la séquence en acides aminés dudit variant gp160 hybride ; et

n) en aval du site et en séquence :

i) quand Y est tel que défini ci-dessus, une troisième région d'ADN codant pour ladite deuxième région de la séquence en acides aminés dudit variant gp160 hybride et

ii) un codon de terminaison de traduction ;

pour obtenir une cassette d'expression comportant ladite unité d'ADN.

10. Un variant gp160 hybride, non-clivable et soluble, ayant une séquence en acides aminés qui comprend :

i) une première région dérivée de la gp160 d'une première souche du virus HIV et située dans cette dernière gp160 de l'acide aminé en position X à l'acide aminé en position Y, X étant un nombre de 1 à 271 et Y étant un nombre de 306 à 482 ;

ii) une deuxième région dérivée d'un variant gp160 soluble, non-clivable de type A ayant pour origine une deuxième souche du virus HIV et située dans ce dernier variant gp160 de l'acide aminé en position Y + 1 à l'acide aminé en position C-terminale, Y étant tel que défini ci-dessus ; et

- 31 -

- iii) quand X est différent de 1, une troisième région dérivée de la gp160 de ladite deuxième souche du virus HIV et située dans cette dernière gp160 de l'acide aminé en position 1 à l'acide aminé en position X - 1, X étant tel que défini ci-dessus.

5 11. Un variant gp160 hybride, non-clivable et soluble selon la revendication 10, ayant une séquence en acides aminés qui comprend :

10 i) une première région dérivée de la gp160 d'une première souche du virus HIV et située dans cette dernière gp160 de l'acide aminé en position X à l'acide aminé en position Y, X étant un nombre de 1 à 271 et Y étant un nombre de 306 à 476 ;

15 ii) une deuxième région dérivée d'un variant gp160 soluble, non-clivable de type A' ayant pour origine une deuxième souche du virus HIV et située dans ce dernier variant gp160 de l'acide aminé en position Y + 1 à l'acide aminé en position C-terminale, Y étant tel que défini ci-dessus ; et

20 iii) quand X est différent de 1, une troisième région dérivée de la gp160 de ladite deuxième souche du virus HIV et située dans cette dernière gp160 de l'acide aminé en position 1 à l'acide aminé en position X - 1, X étant tel que défini ci-dessus.

25 12. Un variant gp160 hybride, non-clivable et soluble selon la revendication 10 ou 11, ayant une séquence en acides aminés qui comprend :

i) une première région dérivée de la gp160 d'une première souche du virus HIV et située dans cette dernière gp160 de l'acide aminé en position X à l'acide aminé en position Y, X étant un nombre de 1 à 271 et Y étant un nombre de 306 à 482 ;

30 ii) une deuxième région dérivée d'un variant gp160 soluble, non-clivable de type A ayant pour origine une deuxième souche du virus HIV et située dans ce dernier variant gp160 de l'acide aminé en position Y + 1 à l'acide aminé en position C-terminale, Y étant tel que défini ci-dessus et ladite deuxième souche étant la souche HIV-1 Bru ; et

35 iii) quand X est différent de 1, une troisième région dérivée de la gp160 de ladite

- 32 -

deuxième souche du virus HIV et située dans cette dernière gp160 de l'acide aminé en position 1 à l'acide aminé en position X - 1, X étant tel que défini ci-dessus.

5 13. Un variant gp160 hybride, non-clivable et soluble selon n'importe laquelle des revendications 10 à 12, ayant une séquence en acides aminés qui comprend :

10 i) une première région dérivée de la gp160 d'une première souche du virus HIV et située dans cette dernière gp160 de l'acide aminé en position X à l'acide aminé en position Y, X étant un nombre de 1 à 271 et Y étant un nombre de 306 à 482 et ladite première souche étant sélectionnée parmi les souches HIV-1 MN, HIV-1 Eli, HIV-1 RF, HIV-1 SF2C et HIV-1 SC;

15 ii) une deuxième région dérivée d'un variant gp160 soluble, non-clivable de type A ayant pour origine une deuxième souche du virus HIV et située dans ce dernier variant gp160 de l'acide aminé en position Y + 1 à l'acide aminé en position C-terminale, Y étant tel que défini ci-dessus ; et

20 iii) quand X est différent de 1, une troisième région dérivée de la gp160 de ladite deuxième souche du virus HIV et située dans cette dernière gp160 de l'acide aminé en position 1 à l'acide aminé en position X - 1, X étant tel que défini ci-dessus.

25 14. Un variant gp160 hybride, non-clivable et soluble selon la revendication 13, qui comprend :

i) une première région dérivée de la gp160 d'une première souche du virus HIV et située dans cette dernière gp160 de l'acide aminé en position X à l'acide aminé en position Y, X étant un nombre de 1 à 271 et Y étant un nombre de 306 à 482 et ladite première souche étant la souche HIV-1 MN;

30 ii) une deuxième région dérivée d'un variant gp160 soluble, non-clivable de type A ayant pour origine une deuxième souche du virus HIV et située dans ce dernier variant gp160 de l'acide aminé en position Y + 1 à l'acide aminé en position C-terminale, Y étant tel que défini ci-dessus ; et

35 iii) quand X est différent de 1, une troisième région dérivée de la gp160 de ladite deuxième souche du virus HIV et située dans cette dernière gp160 de l'acide

- 33 -

aminé en position 1 à l'acide aminé en position X - 1, X étant tel que défini ci-dessus.

15. Un variant gp160 hybride, non-clivable et soluble selon n'importe laquelle des revendications 10 à 14, ayant une séquence en acides aminés qui comprend :

i) une première région dérivée de la gp160 d'une première souche du virus HIV et située dans cette dernière gp160 de l'acide aminé en position X à l'acide aminé en position Y, X étant un nombre de 1 à 271 et Y étant un nombre de 450 à 482 ;

ii) une deuxième région dérivée d'un variant gp160 soluble, non-clivable de type A' ayant pour origine une deuxième souche du virus HIV et située dans ce dernier variant gp160 de l'acide aminé en position Y + 1 à l'acide aminé en position C-terminale, Y étant tel que défini ci-dessus ; et

iii) quand X est différent de 1, une troisième région dérivée de la gp160 de ladite deuxième souche du virus HIV et située dans cette dernière gp160 de l'acide aminé en position 1 à l'acide aminé en position X - 1, X étant tel que défini ci-dessus.

16. Un variant gp160 hybride, non-clivable et soluble selon n'importe laquelle des revendications 10 à 15, ayant une séquence en acides aminés qui comprend :

i) une première région dérivée de la gp160 d'une première souche du virus HIV et située dans cette dernière gp160 de l'acide aminé en position X à l'acide aminé en position Y, X étant un nombre de 1 à 97 et Y étant un nombre de 306 à 482 ;

ii) une deuxième région dérivée d'un variant gp160 soluble, non-clivable de type A' ayant pour origine une deuxième souche du virus HIV et située dans ce dernier variant gp160 de l'acide aminé en position Y + 1 à l'acide aminé en position C-terminale, Y étant tel que défini ci-dessus ; et

iii) quand X est différent de 1, une troisième région dérivée de la gp160 de ladite deuxième souche du virus HIV et située dans cette dernière gp160 de l'acide aminé en position 1 à l'acide aminé en position X - 1, X étant tel que défini ci-dessus.

17. Une cassette d'expression obtenue par une méthode selon n'importe laquelle des revendications 1 à 9.
- 5 18. Une cassette d'expression comportant une unité d'ADN codant soit pour un précurseur d'un variant gp160 selon n'importe laquelle des revendications 10 à 16 soit pour un variant gp160 selon n'importe laquelle des revendications 10 à 16 et les éléments nécessaires à expression de ladite unité d'ADN.
- 10 19. Une cellule transformée par une cassette d'expression selon la revendication 17 ou 18.
- 15 20. Un vecteur viral dans le génome duquel est inséré une cassette d'expression selon la revendication 17 ou 18.
- 20 21. Un procédé pour produire un variant gp160 selon n'importe laquelle des revendications 10 à 16, qui comprend l'acte de cultiver une cellule selon la revendication 19 et l'acte de récolter ledit variant à partir de la culture.
- 25 22. Un procédé pour produire un variant gp160 selon n'importe laquelle des revendications 10 à 16, qui comprend l'acte d'infecter une culture de cellules avec un vecteur viral selon la revendication 20 et l'acte de récolter ledit variant à partir de la culture.
- 30 23. Une composition pharmaceutique destinée à la prévention ou au traitement du Sida, qui comprend, à titre d'agent thérapeutique, un variant gp160 selon n'importe laquelle des revendications 10 à 16.
- 35 24. Une composition pharmaceutique destinée à la prévention ou au traitement du Sida, qui comprend, à titre d'agent thérapeutique, un vecteur viral selon la revendication 20.

1/20

Figure 1

```

-30      -25      -20      -15      -10      -5       1       5       10      15      20      25
:         :         :         :         :         :         :         :         :         :
a) MRVKE---KYQHLWRWGWKWGTMLLGILMICSATEKLWVTVYYGVPVWKEATTTLFCASD
b) MRVKGIRRNYQH-W---WGWGTMLLGILLMICSATEKLWVTVYYGVPVWKEATTTLFCASD
c) MRARGIERNCQN-W---WKWGIMLLGILMTCSAADNLWVTVYYGVPVWKEATTTLFCASD
d) MRVMEMRKNCQHLWKW----GTMLLGMLMICSAAEDLWVTVYYGVPVWKEATTTLFCASE
e) MRVKGSGRNYQHLWRWG----TMLLGILMICSAAEQWLWVTVYYGVPVWKEATTTLFCASD
f) MKVKGTRRNYQHLWRW----GTLLLGMLMICSATEKLWVTVYYGVPVWKEATTTLFCASD
*.....  ..*. *          ..***.***.***.....*****

      30      35      40      45      50      55      60      65      70      75      80      85
      :         :         :         :         :         :         :         :         :         :
a) AKAYDTEVHNVWATHACVPTDPNPQEVVLVNVVTENFNMWKNMVEQMHEDIISLWDQSLK
b) AKAYDTEVHNVWATQACVPTDPNPQEVVLVNVVTENFNMWKNMVEQMHEDIISLWDQSLK
c) AKSYETEAHNIWATHACVPTDPNPQEIAlENVVTENFNMWKNMVEQMHEDIISLWDQSLK
d) AKAYKTEVHNVWAKHACVPTDPNPQEVLLENVVTENFNMWKNMVEQMHEDIISLWDQSLK
e) AKAYDTEVHNIWATHACVPTDPNPQEVVLGNVTENFNMWKNMVEQMHEDIISLWDQSLK
f) ARAYDTEVHNVWATHACVPTDPNPQEVVLGNVTENFNMWKNMVEQMQEDIISLWDQSLK
*...*.***.***.***.....*****. * .....*****.*****.*****

      90      95      100      105      110      115      120      125      130      135      140      145
      :         :         :         :         :         :         :         :         :         :
a) PCVKLTPLCVSLKCTDLGNATNTN-SSNTNSSSGEMMMEKGEIKNCSFNISTSIRGKVQK
b) PCVKLTPLCVTLNCTDLRNTTNTNNSANNNSNSEGTIKGGEMKNCSFNITTSIRDKMOK
c) PCVKLTPLCVTLNCSDELNRNGTMGNVVTTEEKG-----MKNCSFNVTTVLKDKKQK
d) PCVKLTPLCVTLNCTD-----ANLNGTNVTSSSGGTMENGEIKNCSFQVTTSTRDKTQK
e) PCVKLTPLCVTLNCTNLRNDTSTN-ATNTTSSNRGKM-EGGEMTNCSFNITTSIRSKVQK
f) PCVKLTPLCVTLNCTDLGKATNTN-SSNWKEEI-----KGEIKNCSFNITTSIRDKIQK
*****.***.***.  ..  ..  ...          ..*****.***.***.

      150      155      160              165              170      175      180      185      190
      :         :         :         :         :         :         :         :         :
a) EYAFFYKLDIIPID-----NDTTS-----YTLTSCNTSVITQACPKVSFEPIPIHYC
b) EYALLYKLDIVSID-----NDSTS-----YRLISCNTSVITQACPKISFEPIPIHYC
c) VYALFYRLDIVPIDNDS-----STNSTNYRLINCNTSAITQACPKVSFEPIPIHYC
d) KYALFYKLDVVPIDKGNISPKNNTSNTSYGNYTLIHCNSSVITQACPKVSFEPIPIHYC
e) EYALFYKLDVVPID-----N--TS-----YTLINCNTSVITQACPKVSFEPIPIHYC
f) ENALFRNLDVVPID-----NASTTTNYTNYRLIHCNRSVITQACPKVSFEPIPIHYC
*..  .***.***.          .          *  *. **  *.*****.*****.

```

```

195 200 205 210 215 220 225 230 235 240 245 250
: : : : : : : : : : : :
a) APAGFAILKCNNKTFNGTGPCTNVSTVQCTHGIRPVVSTQLLLNGSLAEEEVVIRSANFT
b) APAGFAILKCNDDKFKSGKGSCKNVSTVQCTHGIRPVVSTQLLLNGSLAEEEVVIRSENFT
c) APAGFAILKCRDKKFNGTGPCTNVSTVQCTHGIRPVVSTQLLLNGSLAEEEVVIIRSENLT
d) TPAGFAILKCNDDKFKNGTGPCKNVSTVQCTHGIRPVVSTQLLLNGSLAEEEVVIRSENFT
e) ARW-FAILNCNNKKFNGTGPCTNVSTVQCTHGIRPVVSTHLLLNGSLAEEEVVLRSENFT
f) TPAGFAILKCNNKTFNGKGPCNTNVSTVQCTHGIRPIVSTQLLLNGSLAEEEVVIRSDNFT
.. ****.*..*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*

255 260 265 270 275 280 285 290 295 300 305 310
: : : : : : : : : : : :
a) DNAKTIIVQLNQSVEINCTRPNNNTRKSIRIQRGPGRAFTVIGKI-GNMRQAHCNISRAK
b) DNAKTIIVHLNESVQINCTRPYNKRKRIHI--GPGRAFYTTKNIIGTIRQAHCNISRAK
c) NNAKNIIAHLNESVKITCARPYQNTQRTP--GLGQSLYTT-RSRSIIGQAHCNISRAQ
d) DNVKTIIVQLNASVQINCTRPNNNTRKSIT--KGPGRVYATGQIIGDIRKAHCNLSRAQ
e) DNAKTIIVQLKEAVEINCTRPNNNTTSIHI--GPGRAFYATGDIIGDIRQAHCNISRAK
f) NNAKTIIVQLNESVAINCTRPNNNTRKSIYI--GPGRAFHTTGRIIGDIRKAHCNISRAQ
.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*

315 320 325 330 335 340 345 350 355 360 365 370
: : : : : : : : : : : :
a) WNATLKQIASKLREQFGNNKTII-FKQSSGGDPEIVTHSFNCGGEFFYCNSTQLFNSTWF
b) WNDTLRQIVSKLKEQF-KNKTIV-FNQSSGGDPEIVMHSFNCGGEFFYCNTSPLFNSTWN
c) WSKTLQQVARKLGTLL--NKTIIFKPSSGGDPEITTHSFNCGGEFFYCNTSGLFNSTWN
d) WNNTLKQVVTKLREQF-DNKTIV-FTSSSGGDPEIVLHSFNCGGEFFYCNTTQLFNSTWN
e) WNNTLKQIVIKLRDQF-ENKTII-FNRSSGGDPEIVMHSFNCGGEFFYCNSTQLFSSTWN
f) WNNTLEQIVKKLREQFGNNKTIV-FNQSSGGDPEIVMHSFNCRGEFFYCNTTQLFNN---
.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*

375 380 385 390 395 400 405 410 415 420 425 430
: : : : : : : : : : : :
a) -NSTWSTEGSNNTGSDTITLPCRIKQFINMWQEVGKAMYAPPISGQIRCSSNITGLLLT
b) GNNTWN----NTTGSNNNITLQCKIKQIINMWQEVGKAMYAPPIEGQIRCSSNITGLLLT
c) I-SAWNNITESNNSTNTNITLQCRKQIIM-VAGRKAIYAPPIERNILCSSNITGLLLT
d) S-----TEGSNNTGGNDTITLPCRIKQIVNMWQEVGKAMYAPPISGQIKCISNITGLLLT
e) -----GTEGSNNTGGNDTITLPCRIKEIINMWQEVGKAMYAPPIKGQVKCSSNITGLLLT
f) ---TWRLNHTGCTKGNNDTIIILPCRIKQIINMWQEVGKAMYAPPIGGQISCSSNITGLLLT
.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*

```

3/20

Figure 1 (suite)

```

      435  440      445  450  455  460  465  470  475  480  485
      :    :      :    :    :    :    :    :    :
a) RDGGNNNGS-----EIFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTKAKRRVVQREK
b) RDGGKDTD---TNDTEIFRPGGGDMRDNRSELYKYKVVTIEPLGVAPTKAKRRVVQREK
c) RDGGINN-----STNETFRPGGGDMRDNRSELYKYKVQIEPLGVAPTRAKRRVVEREK
d) RDGGEDT----TNTTEIFRLGGGNMRDNNRSELYKYKVRIEPLGVAPTRAKRRVVQREK
e) RDGGNSKNGSKNENTEIFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTKAKRRVVQREK
f) RDGGTNV----TNDTEVFRPGGGDMRDNRSELYKYKVIKIEPLGIAPTKAKRRVVQREK
      ****      *.*.*****.*****.*****.*****.*****

```

```

      490  495  500  505  510  515  520  525  530  535  540
      :    :    :    :    :    :    :    :    :    :
a) RAVG-IGALFLGFLGAAGSTMGARSMTLTVQARQLLSGIVQQQNNLLRAIEAQQHLLQLT
b) RA--AIGALFLGFLGAAGSTMGAASVTLTVQARLLLSGIVQQQNNLLRAIEAQQHMLQLT
c) RAIG-LGAMFLGFLGAAGSTMGARSVTLTVQARQLMSGIVQQQNNLLRAIEAQQHLLQLT
d) RAVGTIGAMFLGFLGAAGSTMGAGSITLTVQARHLLSGIVQQQNNLLRAIEAQQHLLQLT
e) RAVGTIGAMFLGFLGAAGSTMGATSMTLTVQARLLLSGIVQQQNNLLRAIEAQQHLLQLT
f) RAVGIVGAMFLGFLGAAGSTMGAVSLTLTVQARQLLSGIVQQQNNLLRAIEAQQHLLQLT
      **      *.*.*****.*****.*****.*****.*****.*****

```

```

      545  550  555  560  565  570  575  580  585  590  595  600
      :    :    :    :    :    :    :    :    :    :
a) VWGIKQLQARILAVERYLKDQQLLGIWGCSGKLICTTAVPWNASWSNKSLEQIWNMMTWM
b) VWGIKQLQARVLAVERYLKDQQLLGIWGCSGKLICTTTVPWNASWSNKSLLDIWNMMTWM
c) VWGIKQLQARILAVERYLKDQQLLGIWGCSGKHICTTNVPWNSSWSNRSLEIWNMMTWM
d) VWGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICTTTVPWNASWSNKSLEIWNMMTWM
e) VWGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICTTTVPWNASWSNKSLEIWNMMTWM
f) VWGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICTTAVPWNASWSNKSLEIWNMMTWM
      *****.*****.*****.*****.*****.*****.*****.*****

```

```

      605  610  615  620  625  630  635  640  645  650  655  660
      :    :    :    :    :    :    :    :    :    :
a) EWDREINNYTSLIHSLEESQNQQEKNEQELLELDKWASLWNWFNITNWLWYIKIFIMIV
b) QWEREIDNYTSLIYSLLEKSQTQQEKNEQELLELDKWASLWNWFDITNWLWYIKIFIMIV
c) EWEREIDNYTGLIYSLLEESQTQQEKNEKELLELDKWASLWNWFSITQWLWYIKIFIMII
d) QWEREIDNYTGIIYNLLEESQNQQEKNEQELLELDKWANLWNWFDITQWLWYIRIFIMIV
e) EWEREIDNYTSLIYTLLEESQNQQEKNEQELLELDKWASLWNWFNITNWLWYIKIFIMIV
f) QWEREIDNYTNTIYTLLEESQNQQEKNEQELLELDKWASLWNWFSITNWLWYIKIFIMIV
      *.*.*****.*****.*****.*****.*****.*****.*****

```

a) -----
b) -----
c) CYKGKN
d) -----
e) -----
f) -----

Figure 2

★

6/20

Figure 2 (suite)

```

355 360 365 370 375 380 385 390 395 400 405 410
:   :   :   :   :   :   :   :   :   :   :
a) CGGEFFYCNSTQLFNSTWFNSTWSTEGSNNTGSDTITLPCRIKQFINMWQEVGKAMYAPP
b) CRGEFLYCMT-----WFLN-WIENKTHRNA-----PCHIKQIINTWHKVGGRNVYLPP
   * **.**** *      ** . * .. .... .      **.**** ** *..... **

415 420 425 430 435 440 445 450 455 460 465 470
:   :   :   :   :   :   :   :   :   :   :
a) ISGQIRCSSNITGLLLTRDGGNNNGSEIFRPGGGMDRDNWRSELYKYKVVKIEPLGVAPT
b) REGELSCNSTVTSIIANMDWQNNNQTNITFSAEVAEL---YRLELGDYKLVEITPIGFAPT
   *.....*..... ** *.....*..... . * ** *.....*..... **

475 480 485 490 495 500 505 510 515 520 525 530
:   :   :   :   :   :   :   :   :   :   :
a) KAKRRVVQREKRAVGIGAL-FLGFLGAAGSTMGARSMTLTVQARQLLSGIVQQQNNLLRAI
b) KEKRYSSAHGRHTRGVFVLGFLGFLATAGSAMGAASLTVSAQSRTLLAGIVQQQQQLLDVV
   *.....*.....* . * *.....*.....*.....*.....*.....*.....*.....

535 540 545 550 555 560 565 570 575 580 585 590 595
:   :   :   :   :   :   :   :   :   :   :
a) EAQQHLLQLTVWGIKQLQARILAVERYLKDQQLLGIWGCSGKLICTTAVPWNASWSNKSLEQ
b) KRQQELLRLTVWGTKNLQARVTAIEKYLQDQARLNSWGCAFRQVCHTTVPW----VNDSLAP
   *.....*.....*.....*.....*.....*.....*.....*.....*.....*.....

600 605 610 615 620 625 630 635 640 645 650 655
:   :   :   :   :   :   :   :   :   :   :
a) IWNMTWMEWDREINNYTSLIHSLEESQNQQEKNEQELLELDKWASLWNWFNITNLWYIK
b) DWDNMTWQEWKQVRYLEANISKSLQAQIQQEKMYELQKLNSWDIFGNWFDLTSWVKYIQ
   *.....*.....*.....*.....*.....*.....*.....*.....*.....*.....

660 665 670 675 680 685 690 695 700 705 710
:   :   :   :   :   :   :   :   :   :   :
a) IFIMIVGGLVGLRIVFAVLSIVNRVRQYSPL-----SF--QTHLPTPRGPDREGEIEEGG
b) YGVLIIVAVIALRIVYVQMLSRRLRGYRPFSSPPGYIQQIHKDRGQFANEETEEDGG
   *.....*.....*.....*.....*.....*.....*.....*.....*.....*.....

715 720 725 730 735 740 745 750 755 760 765
:   :   :   :   :   :   :   :   :   :   :
a) ERDRDR-SIRLVNGSLALIWDLLRSLC-LFSYHR--LRDLLIVTRIVELLGRRGWEALK
b) SNGGDRYWPWPIAYIHFLIRQLIRLLTRLYSICRDLLSRSLTLQLIYQNL--RDW--LR
   ... ** . .. **..* * * * * ..*.....*.....*.....*.....*.....

```

Figure 2 (suite')

830
:
a) LERILL
b) AEIALL
* **

830
:
a) LERILL
b) AEIALL
* **

8/20

Figure 3

PstI
 CTGCAGTGACAATGAGAGTGAAGGAGAAATATCAGCACTTGTGGAGATGGGGG 53
 MetArgValLysGluLysTyrGlnHisLeuTrpArgTrpGly

TGGAAATGGGGCACCATGCTCCTTGGGATATTGATGATCTGTAGTGCTACA 104
 TrpLysTrpGlyThrMetLeuLeuGlyIleLeuMetIleCysSerAlaThr
 KpnI

GAAAAATTGTGGGTCACAGTCTATTATGGGGTACCTGTGTGGAAGGAAGCA 155
 GluLysLeuTrpValThrValTyrTyrGlyValProValTrpLysGluAla

ACCACCACTCTATTTTGTGCATCAGATGCTAAAGCATATGATACAGAGGTA 206
 ThrThrThrLeuPheCysAlaSerAspAlaLysAlaTyrAspThrGluVal

CATAATGTTTGGGCCACACATGCCTGTGTACCCACAGACCCCAACCCACAA 257
 HisAsnValTrpAlaThrHisAlaCysValProThrAspProAsnProGln

GAAGTAGTATTGGTAAATGTGACAGAAAAATTTAACATGTGGAAAAATGAC 308
 GluValValLeuValAsnValThrGluAsnPheAsnMetTrpLysAsnAsp

ATGGTAGAACAGATGCATGAGGATATAATCAGTTTATGGGATCAAAGCCTA 359
 MetValGluGlnMetHisGluAspIleIleSerLeuTrpAspGlnSerLeu

AAGCCATGTGTAAAATTAACCCCACTCTGTGTTAGTTTAAAGTGCACTGAT 410
 LysProCysValLysLeuThrProLeuCysValSerLeuLysCysThrAsp

TTGGGGAATGCTACTAATAACCAATAGTAGTAATACCAATAGTAGTAGCGGG 461
 LeuGlyAsnAlaThrAsnThrAsnSerSerAsnThrAsnSerSerSerGly

GAAATGATGATGGAGAAAGGAGAGATAAAAACTGCTCTTTCAATATCAGC 512
 GluMetMetMetGluLysGlyGluIleLysAsnCysSerPheAsnIleSer

ACAAGCATAAGAGGTAAGGTGCAGAAAGAATATGCATTTTTTTATAAACTT 563
 ThrSerIleArgGlyLysValGlnLysGluTyrAlaPhePheTyrLysLeu

GATATAATACCAATAGATAATGATACTACCAGCTATACGTTGACAAGTTGT 614
 AspIleIleProIleAspAsnAspThrThrSerTyrThrLeuThrSerCys

AACACCTCAGTCATTACACAGGCCTGTCCAAAGGTATCCTTTGAGCCAATT 665
 AsnThrSerValIleThrGlnAlaCysProLysValSerPheGluProIle

CCCATACATTATTGTGCCCCGGCTGGTTTTGCGATTCTAAAATGTAATAAT 716
 ProIleHisTyrCysAlaProAlaGlyPheAlaIleLeuLysCysAsnAsn

AAGACGTTCAATGGAACAGGACCATGTACAAATGTCAGCACAGTACAATGT 767
 LysThrPheAsnGlyThrGlyProCysThrAsnValSerThrValGlnCys

ACACATGGAATTAGGCCAGTAGTATCAACTCAACTGCTGTTGAATGGCAGT 818
 ThrHisGlyIleArgProValValSerThrGlnLeuLeuLeuAsnGlySer

CTAGCAGAAGAAGAGGTAGTAATTAGATCTGCCAATTTACAGACAATGCT 869
 LeuAlaGluGluGluValValIleArgSerAlaAsnPheThrAspAsnAla

9/20

Figure 3 (suite)

AAAACCATAATAGTACAGCTGAACCAATCTGTAGAAATTAATTGTACAAGA LysThrIleIleValGlnLeuAsnGlnSerValGluIleAsnCysThrArg	920	
CCCAACAACAATACAAGAAAAAGTATCCGTATCCAGAGGGGACCAGGGAGA ProAsnAsnAsnThrArgLysSerIleArgIleGlnArgGlyProGlyArg	971	
GCATTTGTTACAATAGGAAAAATAGGAAATATGAGACAAGCACATTGTAAC AlaPheValThrIleGlyLysIleGlyAsnMetArgGlnAlaHisCysAsn	1022	
ATTAGTAGAGCAAAATGGAATGCCACTTTAAAACAGATAGCTAGCAAATTA IleSerArgAlaLysTrpAsnAlaThrLeuLysGlnIleAlaSerLysLeu	1073	
AGAGAACAATTTGGAAATAATAAAACAATAATCTTTAAGCAATCCTCAGGA ArgGluGlnPheGlyAsnAsnLysThrIleIlePheLysGlnSerSerGly	1124	
GGGGACCCAGAAATTGTAACGCACAGTTTTAATTGTGGAGGGGAATTTTC GlyAspProGluIleValThrHisSerPheAsnCysGlyGlyGluPhePhe	1175	
TACTGTAATTCAACACAACACTGTTTAATAGTACTTGGTTTAATAGTACTTGG TyrCysAsnSerThrGlnLeuPheAsnSerThrTrpPheAsnSerThrTrp	1226	
AGTACTGAAGGGTCAAATAACACTGAAGGAAGTGACACAATCACACTCCCA SerThrGluGlySerAsnAsnThrGluGlySerAspThrIleThrLeuPro	1277	
TGCAGAATAAAACAATTTATAAACATGTGGCAGGAAGTAGGAAAAGCAATG CysArgIleLysGlnPheIleAsnMetTrpGlnGluValGlyLysAlaMet	1328	
TATGCCCTCCCATCAGCGGACAAATTAGATGTTCATCAAATATTACAGGG TyrAlaProProIleSerGlyGlnIleArgCysSerSerAsnIleThrGly	1379	
		BglII
CTGCTATTAACAAGAGATGGTGGTAATAACAACAATGGGTCCGAGATCTTC LeuLeuLeuThrArgAspGlyGlyAsnAsnAsnAsnGlySerGluIlePhe	1430	
AGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAA ArgProGlyGlyGlyAspMetArgAspAsnTrpArgSerGluLeuTyrLys	1481	
TATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAAGGCACAG TyrLysValValLysIleGluProLeuGlyValAlaProThrLysAlaGln	1532	
AATCACGTGGTGCAGAATGAACACCAAGCAGTGGGAATAGGAGCTTTGTTC AsnHisValValGlnAsnGluHisGlnAlaValGlyIleGlyAlaLeuPhe	1583	
CTTGGGTTCTTGGGAGCAGCAGGAAGCACTATGGGCGCACGGTCAATGACG LeuGlyPheLeuGlyAlaAlaGlySerThrMetGlyAlaArgSerMetThr	1634	
CTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAGCAGCAGAAC LeuThrValGlnAlaArgGlnLeuLeuSerGlyIleValGlnGlnGlnAsn	1685	
AATTTGCTGAGGGCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTC AsnLeuLeuArgAlaIleGluAlaGlnGlnHisLeuLeuGlnLeuThrVal	1736	

10/20

Figure 3 (suite)

TGGGGCATCAAGCAGCTCCAGGCAAGAATCCTGGCTGTGGAAAGATACCTA 1787
TrpGlyIleLysGlnLeuGlnAlaArgIleLeuAlaValGluArgTyrLeu

AAGGATCAACAGCTCCTGGGGATTTGGGGTTGCTCTGGAAAACCTCATTTC 1838
LysAspGlnGlnLeuLeuGlyIleTrpGlyCysSerGlyLysLeuIleCys

ACCACTGCTGTGCCTTGGAAATGCTAGTTGGAGTAATAAATCTCTGGAACAG 1889
ThrThrAlaValProTrpAsnAlaSerTrpSerAsnLysSerLeuGluGln

ATTTGGAATAACATGACCTGGATGGAGTGGGACAGAGAAATTAACAATTAC 1940
IleTrpAsnAsnMetThrTrpMetGluTrpAspArgGluIleAsnAsnTyr

ACAAGCTTAATACATTCTTAATTGAAGAATCGCAAAACCAGCAAGAAAAG 1991
ThrSerLeuIleHisSerLeuIleGluGluSerGlnAsnGlnGlnGluLys

AATGAACAAGAATTATTGGAATTAGATAAATGGGCAAGTTTGTGGAATTGG 2042
AsnGluGlnGluLeuLeuGluLeuAspLysTrpAlaSerLeuTrpAsnTrp

TTTAACATAACAAATTGGCTGTGGTATATAAAAAATAGAGTTAGGCAGGGA 2093
PheAsnIleThrAsnTrpLeuTrpTyrIleLysAsnArgValArgGlnGly

TATTCACCATTATCGTTTCAGACCCACCTCCCAACCCCGAGGGGACCCGAC 2144
TyrSerProLeuSerPheGlnThrHisLeuProThrProArgGlyProAsp

AGGCCCCAAGGAATAGAAGAAGAAGGTGGAGAGAGAGACAGAGACAGATCC 2195
ArgProGluGlyIleGluGluGluGlyGlyGluArgAspArgAspArgSer

ATTCGATTAGTGAACGGATCCTTAGCACTTATCTGGGACGATCTGCGGAGC 2246
IleArgLeuValAsnGlySerLeuAlaLeuIleTrpAspAspLeuArgSer

CTGTGCCTCTTCAGCTACCACCGCTTGAGAGACTTACTCTTGATTGTAACG 2297
LeuCysLeuPheSerTyrHisArgLeuArgAspLeuLeuLeuIleValThr

AGGATTGTGGAACCTTCTGGGACGCAGGGGGTGGGAAGCCCTCAAATATTGG 2348
ArgIleValGluLeuLeuGlyArgArgGlyTrpGluAlaLeuLysTyrTrp

TGGAATCTCCTACAGTATTGGAGTCAGGAATAAAGAATAGTGCTGTTAGC 2399
TrpAsnLeuLeuGlnTyrTrpSerGlnGluLeuLysAsnSerAlaValSer

TTGCTCAATGCCACAGCCATAGCAGTAGCTGAGGGGACAGATAGGGTTATA 2450
LeuLeuAsnAlaThrAlaIleAlaValAlaGluGlyThrAspArgValIle

GAAGTAGTACAAGGAGCTTGTAGAGCTATTCGCCACATACCTAGAAGAATA 2501
GluValValGlnGlyAlaCysArgAlaIleArgHisIleProArgArgIle

AGACAGGGCTTGGAAAGGATTTTGCTATAAGATGGGTGGCAAGTGGTCAAA 2552
ArgGlnGlyLeuGluArgIleLeuLeuTerAspGlyTrpGlnValValLys

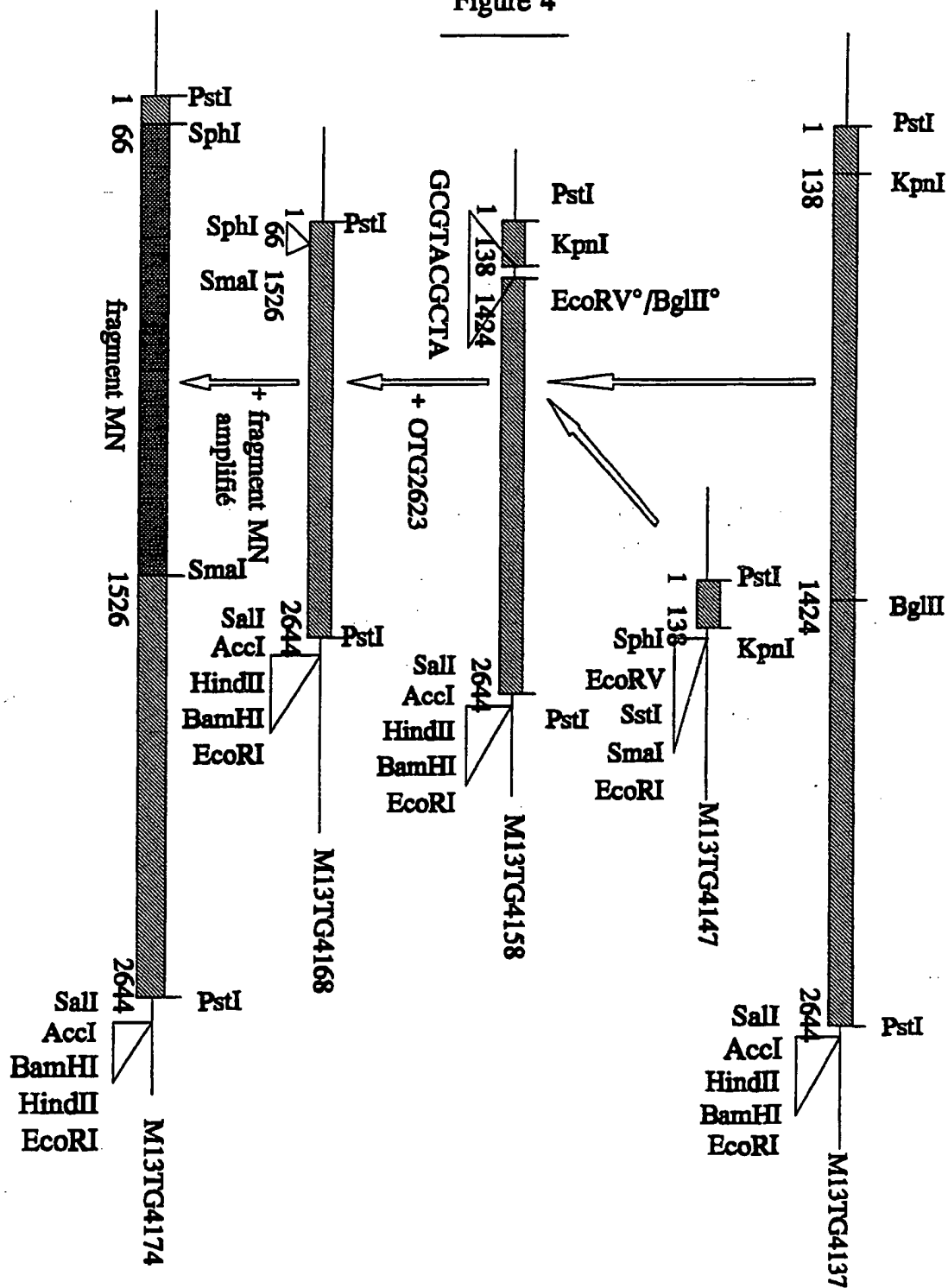
11/20

Figure 3 (suite)

AAGTAGTGTGGTTGGATGGCCTACTGTAAGGGAAAGAATGAGACGAGCTGA 2603
LysTerCysGlyTrpMetAlaTyrCysLysGlyLysAsnGluThrSer***

PstI
GCCAGCAGCAGATGGGGTGGGAGCAGCATCTCGACCTGCAG 2644

12/20
Figure 4



13/20

Figure 5

ATGAGAGTGAAGGGGATCAGGAGGAATTATCAGCACTGGTGGGGATGGGGC	51	GGC
MetArgValLysGlyIleArgArgAsnTyrGlnHisTrpTrpGlyTrpGly		
OTG2624		
AGCATGCTCCTTGGGATATTGATGATCTG		
ACGATGCTCCTTGGGTATTAAATGATCTGTAGTGCTACAGAAAAATTGTGG	102	
ThrMetLeuLeuGlyLeuLeuMetIleCysSerAlaThrGluLysLeuTrp		
GTCACAGTCTATTATGGGGTACCTGTGTGGAAAGAAGCAACCACCACTCTA	153	
ValThrValTyrTyrGlyValProValTrpLysGluAlaThrThrThrLeu		
TTTTGTGCATCAGATGCTAAAGCATATGATACAGAGGTACATAATGTTTGG	204	
PheCysAlaSerAspAlaLysAlaTyrAspThrGluValHisAsnValTrp		
GCCACACAAGCCTGTGTACCCACAGACCCCAACCCACAAGAAGTAGAATTG	255	
AlaThrGlnAlaCysValProThrAspProAsnProGlnGluValGluLeu		
GTAAATGTGACAGAAAAATTTTAACATGTGGAAAAATAACATGGTAGAACAG	306	
ValAsnValThrGluAsnPheAsnMetTrpLysAsnAsnMetValGluGln		
ATGCATGAGGATATAATCAGTTTATGGGATCAAAGCCTAAAGCCATGTGTA	357	
MetHisGluAspIleIleSerLeuTrpAspGlnSerLeuLysProCysVal		
AAATTAACCCCACTCTGTGTTACTTTAAATTGCACTGATTTGAGGAATACT	408	
LysLeuThrProLeuCysValThrLeuAsnCysThrAspLeuArgAsnThr		
ACTAATACCAATAATAGTACTGCTAATAACAATAGTAATAGCGAGGGAACA	459	
ThrAsnThrAsnAsnSerThrAlaAsnAsnAsnSerAsnSerGluGlyThr		
ATAAAGGGAGGAGAAATGAAAACTGCTCTTTCAATATCACCACAAGCATA	510	
IleLysGlyGlyGluMetLysAsnCysSerPheAsnIleThrThrSerIle		
AGAGATAAGATGCAGAAAGAATATGCACTTCTTTATAAACTTGATATAGTA	561	
ArgAspLysMetGlnLysGluTyrAlaLeuLeuTyrLysLeuAspIleVal		
TCAATAGATAATGATAGTACCAGCTATAGGTTGATAAGTTGTAATACCTCA	612	
SerIleAspAsnAspSerThrSerTyrArgLeuIleSerCysAsnThrSer		
GTCATTACACAAGCTTGTCCAAAGATATCCTTTGAGCCAATTCCCATACAC	663	
ValIleThrGlnAlaCysProLysIleSerPheGluProIleProIleHis		
TATTGTGCCCCGGCTGGTTTTTCGATTCTAAAATGTAACGATAAAAAGTTC	714	
TyrCysAlaProAlaGlyPheAlaIleLeuLysCysAsnAspLysLysPhe		
AGTGGAAAAGGATCATGTAAAAATGTCAGCACAGTACAATGTACACATGGA	765	
SerGlyLysGlySerCysLysAsnValSerThrValGlnCysThrHisGly		
ATTAGGCCAGTAGTATCAACTCAACTGCTGTAAATGGCAGTCTAGCAGAA	816	
IleArgProValValSerThrGlnLeuLeuLeuAsnGlySerLeuAlaGlu		

14/20

Figure 5 (suite)

GAAGAGGTAGTAATTAGATCTGAGAATTTCACTGATAATGCTAAAACCATC GluGluValValIleArgSerGluAsnPheThrAspAsnAlaLysThrIle	867
ATAGTACATCTGAATGAATCTGTACAAATTAATTGTACAAGACCCAACTAC IleValHisLeuAsnGluSerValGlnIleAsnCysThrArgProAsnTyr	918
AATAAAAGAAAAAGGATACATATAGGACCAGGGAGAGCATTTTATACAACA AsnLysArgLysArgIleHisIleGlyProGlyArgAlaPheTyrThrThr	969
AAAAATATAATAGGAACTATAAGACAAGCACATTGTAACATTAGTAGAGCA LysAsnIleIleGlyThrIleArgGlnAlaHisCysAsnIleSerArgAla	1020
AAATGGAATGACACTTTAAGACAGATAGTTAGCAAATTTAAAGAACAATTT LysTrpAsnAspThrLeuArgGlnIleValSerLysLeuLysGluGlnPhe	1071
AAGAATAAAACAATAGTCCTTTAATCAATCCTCAGGAGGGGACCCAGAAATT LysAsnLysThrIleValPheAsnGlnSerSerGlyGlyAspProGluIle	1122
GTAATGCACAGTTTTAATTGTGGAGGGGAATTTTTCTACTGTAATACATCA ValMetHisSerPheAsnCysGlyGlyGluPhePheTyrCysAsnThrSer	1173
CCACTGTTTAATAGTACTTGGAATGGTAATAATACTTGGAATAATACTACA ProLeuPheAsnSerThrTrpAsnGlyAsnAsnThrTrpAsnAsnThrThr	1224
GGGTCAAATAACAATATCACACTTCAATGCAAAATAAAACAAATTATAAAC GlySerAsnAsnAsnIleThrLeuGlnCysLysIleLysGlnIleIleAsn	1275
ATGTGGCAGGAAGTAGGAAAAGCAATGTATGCCCTCCCATTGAAGGACAA MetTrpGlnGluValGlyLysAlaMetTyrAlaProProIleGluGlyGln	1326
ATTAGATGTTTCATCAAATATTACAGGGCTACTATTAACAAGAGATGGTGGT IleArgCysSerSerAsnIleThrGlyLeuLeuLeuThrArgAspGlyGly	1377
AAGGACACGGACACGAACGACACCGAGATCTTCAGACCTGGAGGAGGAGAT LysAspThrAspThrAsnAspThrGluIlePheArgProGlyGlyGlyAsp	1428
ATGAGGGACAATTGGAGAAGTGAATTATATAAAATATAAAGTAGTAACAATT MetArgAspAsnTrpArgSerGluLeuTyrLysTyrLysValValThrIle	1479
OTG2625	
CTTGGTAATCCTCATCGTGGGTGGGCCCCGTTTC GAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTGCAGAGA GluProLeuGlyValAlaProThrLysAlaLysArgArgValValGlnArg	1530
GAAAAAAGA GluLysArg	1539

15/20

Figure 6

ATGAGAGCGAGGGGGATAGAGAGAAATTGTCAAAACTGGTGGAAATGGGGC 51
MetArgAlaArgGlyIleGluArgAsnCysGlnAsnTrpTrpLysTrpGly

OTG2624

AGCATGCTCCTTGGGATATTGATGATCTG
ATCATGCTCCTTGGGATATTGATGACCTGTAGTGCTGCAGACAATCTGTGG 102
IleMetLeuLeuGlyIleLeuMetThrCysSerAlaAlaAspAsnLeuTrp

GTCACAGTTTATTATGGGGTGCCTGTATGGAAGGAAGCAACCACCACTCTA 153
ValThrValTyrTyrGlyValProValTrpLysGluAlaThrThrThrLeu

TTTTGTGCATCAGATGCTAAATCATATGAAACAGAGGCACATAATATCTGG 204
PheCysAlaSerAspAlaLysSerTyrGluThrGluAlaHisAsnIleTrp

GCCACACATGCCTGTGTACCCACGGACCCCAACCCACAAGAAATAGCACTG 255
AlaThrHisAlaCysValProThrAspProAsnProGlnGluIleAlaLeu

GAAAATGTGACAGAAAACCTTTAACATGTGAAAAATAACATGGTGGAAACAG 306
GluAsnValThrGluAsnPheAsnMetTrpLysAsnAsnMetValGluGln

ATGCATGAGGATATAATCAGTTTATGGGATCAAAGCCTAAAACCATGTGTA 357
MetHisGluAspIleIleSerLeuTrpAspGlnSerLeuLysProCysVal

AAATTAACCCCACTCTGTGTCACTTTAAACTGTAGTGATGAATTGAGGAAC 408
LysLeuThrProLeuCysValThrLeuAsnCysSerAspGluLeuArgAsn

AATGGCACTATGGGGAACAATGTCACTACAGAGGAGAAAGGAATGAAAAAC 459
AsnGlyThrMetGlyAsnAsnValThrThrGluGluLysGlyMetLysAsn

TGCTCTTTCAATGTAACCACAGTACTAAAAGATAAGAAGCAGCAAGTATAT 510
CysSerPheAsnValThrThrValLeuLysAspLysLysGlnGlnValTyr

GCACTTTTTTATAGACTTGATATAGTACCAATAGACAATGATAGTAGTACC 561
AlaLeuPheTyrArgLeuAspIleValProIleAspAsnAspSerSerThr

AATAGTACCAATTATAGGTTAATAAATTGTAATACCTCAGCCATTACACAG 612
AsnSerThrAsnTyrArgLeuIleAsnCysAsnThrSerAlaIleThrGln

GCTTGTCCAAAGGTATCCTTTGAGCCAATTCCCATACATTATTGTGCCCCA 663
AlaCysProLysValSerPheGluProIleProIleHisTyrCysAlaPro

GCTGGTTTTGCGATTCTAAAGTGTAGAGATAAGAAGTTCAATGGAACAGGC 714
AlaGlyPheAlaIleLeuLysCysArgAspLysLysPheAsnGlyThrGly

CCATGCACAAATGTCAGCACAGTACAATGTACACATGGAATTAGGCCAGTG 765
ProCysThrAsnValSerThrValGlnCysThrHisGlyIleArgProVal

GTGTCAACTCAACTGCTGTTGAATGGCAGTCTAGCAGAAGAAGAGGTCATA 816
ValSerThrGlnLeuLeuLeuAsnGlySerLeuAlaGluGluGluValIle

16/20

Figure 6 (suite)

ATTAGATCCGAAAATCTCACAAACAATGCTAAAAACATAATAGCACATCTT IleArgSerGluAsnLeuThrAsnAsnAlaLysAsnIleIleAlaHisLeu	867
AATGAATCTGTAAAAATTACCTGTGCAAGGCCCTATCAAAATACAAGACAA AsnGluSerValLysIleThrCysAlaArgProTyrGlnAsnThrArgGln	918
AGAACACCTATAGGACTAGGGCAATCACTCTATACTACAAGATCAAGATCA ArgThrProIleGlyLeuGlyGlnSerLeuTyrThrThrArgSerArgSer	969
ATAATAGGACAAGCACATTGTAATATTAGTAGAGCACAATGGAGTAAAACT IleIleGlyGlnAlaHisCysAsnIleSerArgAlaGlnTrpSerLysThr	1020
TTACAACAAGTAGCTAGAAAATTAGGAACCCCTTCTTAACAAAACAATAATA LeuGlnGlnValAlaArgLysLeuGlyThrLeuLeuAsnLysThrIleIle	1071
AAGTTTAAACCATCCTCAGGAGGGGACCCAGAAATTACAACACACAGTTTT LysPheLysProSerSerGlyGlyAspProGluIleThrThrHisSerPhe	1122
AATTGTGGAGGGGAATTCTTCTACTGTAATACATCAGGACTGTTTAATAGT AsnCysGlyGlyGluPhePheTyrCysAsnThrSerGlyLeuPheAsnSer	1173
ACATGGAATATTAGTGCATGGAATAATATTACAGAGTCAAATAATAGCACA ThrTrpAsnIleSerAlaTrpAsnAsnIleThrGluSerAsnAsnSerThr	1224
AACACAAACATCACACTCCAATGCAGAATAAAACAAATTATAAAGATGGTG AsnThrAsnIleThrLeuGlnCysArgIleLysGlnIleIleLysMetVal	1275
GCAGGCAGGAAAGCAATATATGCCCTCCTATCGAAAGAAACATTCTATGT AlaGlyArgLysAlaIleTyrAlaProProIleGluArgAsnIleLeuCys	1326
TCATCAAATATTACAGGGCTACTATTGACAAGAGATGGTGGTATAAATAAT SerSerAsnIleThrGlyLeuLeuLeuThrArgAspGlyGlyIleAsnAsn	1377
AGTACTAACGAGACCTTTAGACCTGGAGGAGGAGATATGAGGGACAATTGG SerThrAsnGluThrPheArgProGlyGlyGlyAspMetArgAspAsnTrp	1428
CTTGTAATCCTCAT AGAAGTGAATTATATAAATATAAGGTAGTACAAATTGAACCACTAGGAGTA ArgSerGluLeuTyrLysTyrLysValValGlnIleGluProLeuGlyVal	1479
OTG2625 CGTGGGTGGGCCCCGTTC GCACCCACCAGGGCAAAGAGAAGAGTGGTGGAAAGAGAAAAAAGA AlaProThrArgAlaLysArgArgValValGluArgGluLysArg	1524

17/20

Figure 7

ATGAGAGTGATGGAGATGAGGAAGAATTGTCAGCACTTGTGGAAATGGGGC MetArgValMetGluMetArgLysAsnCysGlnHisLeuTrpLysTrpGly	GGC 51
OTG2624 AGCATGCTCCTTGGGATATTGATGATCTG ACCATGCTCCTTGGGATGTTGATGATCTGTAGTGCTGCAGAGGACTTGTGG ThrMetLeuLeuGlyMetLeuMetIleCysSerAlaAlaGluAspLeuTrp	102
GTCACAGTCTATTATGGGGTACCTGTGTGGAAAGAAGCAACCACCACTCTA ValThrValTyrTyrGlyValProValTrpLysGluAlaThrThrThrLeu	153
TTTTGTGCATCAGAAGCTAAAGCATATAAAACAGAGGTACATAATGTCTGG PheCysAlaSerGluAlaLysAlaTyrLysThrGluValHisAsnValTrp	204
GCCAAACATGCTTGTGTACCTACAGACCCCAACCCACAAGAAGTACTATTG AlaLysHisAlaCysValProThrAspProAsnProGlnGluValLeuLeu	255
GAAAATGTGACAGAAAATTTTAACATGTGGAAAAATAACATGGTAGAACAG GluAsnValThrGluAsnPheAsnMetTrpLysAsnAsnMetValGluGln	306
ATGCATGAGGATATAATCAGTTTATGGGATCAAAGCCTAAAGCCATGTGTA MetHisGluAspIleIleSerLeuTrpAspGlnSerLeuLysProCysVal	357
AAATTAACCCCACTCTGTGTTACTTTAAATTGCACTGATGCTAACTTGAAT LysLeuThrProLeuCysValThrLeuAsnCysThrAspAlaAsnLeuAsn	408
GGTACTAATGTCACCTAGTAGTAGCGGGGAACAATGATGGAGAACGGAGAA GlyThrAsnValThrSerSerSerGlyGlyThrMetMetGluAsnGlyGlu	459
ATAAAAACTGCTCTTTCCAAGTTACCACAAGTAGAAGAGATAAGACGCAG IleLysAsnCysSerPheGlnValThrThrSerArgArgAspLysThrGln	510
AAAAAATATGCACTTTTTTATAAACTTGATGTGGTACCAATAGAGAAGGGT LysLysTyrAlaLeuPheTyrLysLeuAspValValProIleGluLysGly	561
AATATTAGCCCTAAGAATAATACTAGCAATAATACTAGCTATGGTAACTAT AsnIleSerProLysAsnAsnThrSerAsnAsnThrSerTyrGlyAsnTyr	612
ACATTGATACATTGTAATTCCTCAGTCATTACACAGGCCTGTCCAAAGGTA ThrLeuIleHisCysAsnSerSerValIleThrGlnAlaCysProLysVal	663
TCCTTTGAGCCAATTCCCATACATTATTGCACCCCGGCTGGTTTTGCGATT SerPheGluProIleProIleHisTyrCysThrProAlaGlyPheAlaIle	714
CTAAAGTGTAATGATAAGAAGTTCAATGGAACAGGACCATGTAAAAATGTC LeuLysCysAsnAspLysLysPheAsnGlyThrGlyProCysLysAsnVal	765
AGCACAGTACAATGTACACATGGAATTAGGCCAGTAGTGTCAACTCAACTG SerThrValGlnCysThrHisGlyIleArgProValValSerThrGlnLeu	816

18/20

Figure 7 (suite)

CTGTTAAATGGCAGTCTAGCAGAAGAAGAGGTAGTAATTAGATCTGAAAAT LeuLeuAsnGlySerLeuAlaGluGluGluValValIleArgSerGluAsn	867
TTCACGGACAATGTTAAAACCATAATAGTACAGCTGAATGCATCTGTACAA PheThrAspAsnValLysThrIleIleValGlnLeuAsnAlaSerValGln	918
ATTAATTGTACAAGACCCAACAACAATACAAGAAAAAGTATAACTAAGGGA IleAsnCysThrArgProAsnAsnAsnThrArgLysSerIleThrLysGly	969
CCAGGGAGAGTAATTTATGCAACAGGACAAATAATAGGAGATATAAGAAAA ProGlyArgValIleTyrAlaThrGlyGlnIleIleGlyAspIleArgLys	1020
GCACATTGTAACCTTAGTAGAGCACAATGGAATAACACTTTAAAACAGGTA AlaHisCysAsnLeuSerArgAlaGlnTrpAsnAsnThrLeuLysGlnVal	1071
GTTACAAAATTAAGAGAACAATTTGACAATAAAACAATAGTCTTTACGTCA ValThrLysLeuArgGluGlnPheAspAsnLysThrIleValPheThrSer	1122
TCCTCAGGAGGGGACCCAGAAATTGTACTTCACAGTTTTAATTGTGGAGGG SerSerGlyGlyAspProGluIleValLeuHisSerPheAsnCysGlyGly	1173
GAATTTTCTACTGTAATACAACACAACCTGTTTAATAGTACTTGAATAGT GluPhePheTyrCysAsnThrThrGlnLeuPheAsnSerThrTrpAsnSer	1224
ACTGAAGGGTCAAATAACACTGGAGGAAATGACACAATCACACTCCCATGC ThrGluGlySerAsnAsnThrGlyGlyAsnAspThrIleThrLeuProCys	1275
AGAATAAAACAAATTGTAAACATGTGGCAGGAAGTAGGAAAAGCAATGTAT ArgIleLysGlnIleValAsnMetTrpGlnGluValGlyLysAlaMetTyr	1326
GCCCCCTCCCATCAGTGGACAAATTAAATGTATATCAAATATTACAGGGCTA AlaProProIleSerGlyGlnIleLysCysIleSerAsnIleThrGlyLeu	1377
CTATTAACAAGAGATGGGGGTGAAGATACAATAACTACTACAGAGATCTTC LeuLeuThrArgAspGlyGlyGluAspThrThrAsnThrThrGluIlePhe	1428
AGACTTGGAGGAGGAAATATGAGGGACAATTGGAGAAGTGAATTATATAAA ArgLeuGlyGlyGlyAsnMetArgAspAsnTrpArgSerGluLeuTyrLys	1479
OTG2625	
CTTGGTAATCCTCATCGTGGGTGGGCCCCGTTTG TATAAAGTGGTAAGAATTGAGCCATTAGGAGTGGCACCCACTAGGGCAAAG TyrLysValValArgIleGluProLeuGlyValAlaProThrArgAlaLys	1530
AGAAGAGTGGTGCAAAGAGAAAAAAGA ArgArgValValGlnArgGluLysArg	1557

19/20

Figure 8

ATGAAAGTGAAGGGGACCAGGAGGAATTATCAGCACTTGTGGAGATGGGGC MetLysValLysGlyThrArgArgAsnTyrGlnHisLeuTrpArgTrpGly	51	GGC
OTG2624		
AGCATGCTCCTTGGGATATTGATGATCTG ACCTTGCTCCTTGGGATGTTGATGATCTGTAGTGCTACAGAAAAATTGTGG ThrLeuLeuLeuGlyMetLeuMetIleCysSerAlaThrGluLysLeuTrp	102	
GTCACAGTTTATTATGGAGTACCTGTGTGGAAAGAAGCAACTACCACTCTA ValThrValTyrTyrGlyValProValTrpLysGluAlaThrThrThrLeu	153	
TTTTGTGCATCAGATGCTAGAGCATATGATACAGAGGTACATAATGTTTGG PheCysAlaSerAspAlaArgAlaTyrAspThrGluValHisAsnValTrp	204	
GCCACACATGCCTGTGTACCCACAGACCCCAACCCACAAGAAGTAGTATTG AlaThrHisAlaCysValProThrAspProAsnProGlnGluValValLeu	255	
GGAAATGTGACAGAAAATTTTAACATGTGGAAAAATAACATGGTAGAACAG GlyAsnValThrGluAsnPheAsnMetTrpLysAsnAsnMetValGluGln	306	
ATGCAGGAGGATATAATCAGTTTATGGGATCAAAGCCTAAAGCCATGTGTA MetGlnGluAspIleIleSerLeuTrpAspGlnSerLeuLysProCysVal	357	
AAATTAACCCCACTCTGTGTTACTTTAAATTGCACTGATTTGGGGAAGGCT LysLeuThrProLeuCysValThrLeuAsnCysThrAspLeuGlyLysAla	408	
ACTAATACCAATAGTAGTAATTGGAAAGAAGAAATAAAAGGAGAAATAAAA ThrAsnThrAsnSerSerAsnTrpLysGluGluIleLysGlyGluIleLys	459	
AACTGCTCTTTCAATATCACCACAAGCATAAGAGATAAGATTCAGAAAGAA AsnCysSerPheAsnIleThrThrSerIleArgAspLysIleGlnLysGlu	510	
AATGCACTTTTTCGTAACCTTGATGTAGTACCAATAGATAATGCTAGTACT AsnAlaLeuPheArgAsnLeuAspValValProIleAspAsnAlaSerThr	561	
ACTACCAACTATAACCAACTATAGGTTGATACATTGTAACAGATCAGTCATT ThrThrAsnTyrThrAsnTyrArgLeuIleHisCysAsnArgSerValIle	612	
ACACAGGCCTGTCCAAAGGTATCATTTGAGCCAATTCCCATACATTATTGT ThrGlnAlaCysProLysValSerPheGluProIleProIleHisTyrCys	663	
ACCCCGGCTGGTTTTGCGATTCTAAAGTGTAATAATAAACGTTCAATGGA ThrProAlaGlyPheAlaIleLeuLysCysAsnAsnLysThrPheAsnGly	714	
AAAGGACCATGTACAAATGTCAGCACAGTACAATGTACACATGGAATTAGG LysGlyProCysThrAsnValSerThrValGlnCysThrHisGlyIleArg	765	
CCAATAGTGTCAACTCAACTGCTGTAAATGGCAGTCTAGCAGAAGAAGAG ProIleValSerThrGlnLeuLeuLeuAsnGlySerLeuAlaGluGluGlu	816	

20/20

Figure 8 (suite)

GTAGTAATTAGATCTGACAATTTACGAACAATGCTAAAACCATAATAGTA ValValIleArgSerAspAsnPheThrAsnAsnAlaLysThrIleIleVal	867
CAGCTGAATGAATCTGTAGCAATTAAGTGTACAAGACCCAACAACAATACA GlnLeuAsnGluSerValAlaIleAsnCysThrArgProAsnAsnAsnThr	918
AGAAAAAGTATCTATATAGGACCAGGGAGAGCATTTCATACAACAGGAAGA ArgLysSerIleTyrIleGlyProGlyArgAlaPheHisThrThrGlyArg	969
ATAATAGGAGATATAAGAAAAGCACATTGTAACATTAGTAGAGCACAATGG IleIleGlyAspIleArgLysAlaHisCysAsnIleSerArgAlaGlnTrp	1020
AATAACACTTTAGAACAGATAGTTAAAAAATTAAGAGAACAGTTTGGAAT AsnAsnThrLeuGluGlnIleValLysLysLeuArgGluGlnPheGlyAsn	1071
AATAAAACAATAGTCTTTAATCAATCCTCAGGAGGGGACCCAGAAATTGTA AsnLysThrIleValPheAsnGlnSerSerGlyGlyAspProGluIleVal	1122
ATGCACAGTTTTAATTGTAGAGGGGAATTTTTCTACTGTAATACAACACAA MetHisSerPheAsnCysArgGlyGluPhePheTyrCysAsnThrThrGln	1173
CTGTTTAATAATACATGGAGGTTAAATCACACTGAAGGAACTAAAGGAAAT LeuPheAsnAsnThrTrpArgLeuAsnHisThrGluGlyThrLysGlyAsn	1224
GACACAATCATACTCCCATGTAGAATAAAACAAATTATAAACATGTGGCAG AspThrIleIleLeuProCysArgIleLysGlnIleIleAsnMetTrpGln	1275
GAAGTAGGAAAAGCAATGTATGCCCTCCCATTTGGAGGACAAATTAGTTGT GluValGlyLysAlaMetTyrAlaProProIleGlyGlyGlnIleSerCys	1326
TCATCAAATATTACAGGGCTGCTATTAACAAGAGATGGTGGTACAAATGTA SerSerAsnIleThrGlyLeuLeuLeuThrArgAspGlyGlyThrAsnVal	1377
ACTAATGACACCGAGGTCTTCAGACCTGGAGGAGGAGATATGAGGGACAAT ThrAsnAspThrGluValPheArgProGlyGlyGlyAspMetArgAspAsn	1428
CTTGGAATCCT TGGAGAAGTGAATTATATAAATATAAAGTAATAAAAATTGAACCATTAGGA TrpArgSerGluLeuTyrLysTyrLysValIleLysIleGluProLeuGly	1479
OTG2625 CATCGTGGGTGGGCCCCGTTTC ATAGCACCCACCAAGGCAAAGAGAAGAGTGGTGCAGAGAGAAAAAAGA IleAlaProThrLysAlaLysArgArgValValGlnArgGluLysArg	1527

INTERNATIONAL SEARCH REPORT

International application No.
PCT/FR 92/00394

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl.⁵ C12N15/48; C12N15/62; C12N15/86; C12N5/10; A61K39/21

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl.⁵ C12N ; C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	EP,A,0314534 (TRANSGENE S.A. ET AL.) 3 May 1989, the whole document	1, 10, 17-24
Y	WO,A,9102544 (INSTITUT PASTEUR ET AL.) 7 March 1991, the whole document	1, 10, 17-24
A	EP,A,0370458 (ABBOT LABORATORIES) 30 May 1990 the whole document	1, 10, 17-19, 21, 23
Y	EP,A,0245136 (TRANSGENE S.A. ET AL.) 11 November 1987, cited in the application the whole document	1-3, 10, 17-24
Y	WO,A,9012880 (APPLIED BIOTECHNOLOGY, INC.) 1 November 1990, whole document	1-3, 10, 17-24

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

- * Special categories of cited documents:
- "A" document defining the general state of the art which is not considered to be of particular relevance
 - "E" earlier document but published on or after the international filing date
 - "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
 - "O" document referring to an oral disclosure, use, exhibition or other means
 - "P" document published prior to the international filing date but later than the priority date claimed
 - "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
 - "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
 - "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
 - "&" document member of the same patent family

Date of the actual completion of the international search 06 August 1992 (06.08.92)	Date of mailing of the international search report 07 September 1992 (07.09.92)
Name and mailing address of the ISA/ EUROPEAN PATENT OFFICE	Authorized officer
Facsimile No.	Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/FR 92/00394

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	<p>WO,A,9105864 (CONNAUGHT LABORATORIES LIMITED) 2 May 1991</p> <p>whole document</p>	<p>1-6, 10-14, 17-19, 21,23</p>
P,X	<p>WO,A,9107664 (CAMBRIDGE BIOSCIENCE CORPORATION) 30 May 1991</p> <p>whole document</p>	<p>1,10, 17-19, 21,23</p>
P,X	<p>WO,A, 9115238 (GENENTECH, INC.) 17 October 1991</p> <p>whole document</p>	<p>1,10, 17-19, 21,23</p>

**ANNEX TO THE INTERNATIONAL SEARCH REPORT
ON INTERNATIONAL PATENT APPLICATION NO. FR 9200394
SA 59743**

This annex lists the patent family members relating to the patent documents cited in the above-mentioned international search report.
The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information. 06/08/92

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP-A-0314534	03-05-89	FR-A- 2620030	10-03-89
		AU-A- 2190288	09-03-89
		JP-A- 2000448	05-01-90
WO-A-9102544	07-03-91	FR-A- 2650954	22-02-91
		EP-A- 0439601	07-08-91
EP-A-0370458	30-05-90	AU-A- 4545889	05-07-90
		CA-A- 2003383	23-05-90
		JP-A- 2273187	07-11-90
EP-A-0245136	11-11-87	FR-A- 2596771	09-10-87
		FR-A, B 2606029	06-05-88
		AU-B- 604696	03-01-91
		AU-A- 7234987	09-11-87
		WO-A- 8706260	22-10-87
		JP-T- 1500161	26-01-89
WO-A-9012880	01-11-90	EP-A- 0469089	05-02-92
WO-A-9105864	02-05-91	EP-A- 0495811	29-07-92
WO-A-9107664	30-05-91	EP-A- 0453563	30-10-91
WO-A-9115238	17-10-91	AU-A- 7584791	30-10-91

RAPPORT DE RECHERCHE INTERNATIONALE

Demande Internationale No

PCT/FR 92/00394

I. CLASSEMENT DE L'INVENTION (si plusieurs symboles de classification sont applicables, les indiquer tous) ⁷		
Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB		
CIB 5 C12N15/48; A61K39/21	C12N15/62;	C12N15/86; C12N5/10
II. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE		
Documentation minimale consultée ⁸		
Système de classification	Symboles de classification	
CIB 5	C12N ; C07K	
Documentation consultée autre que la documentation minimale dans la mesure où de tels documents font partie des domaines sur lesquels la recherche a porté		
III. DOCUMENTS CONSIDERES COMME PERTINENTS ¹⁰		
Catégorie ⁹	Identification des documents cités, avec indication, si nécessaire ¹² des passages pertinents ¹³	No. des revendications visées ¹⁴
Y	EP,A,0 314 534 (TRANSGENE S.A. ET AL.) 3 Mai 1989 * Document entier *	1, 10, 17-24
Y	WO,A,9 102 544 (INSTITUT PASTEUR ET AL.) 7 Mars 1991 * Document entier *	1, 10, 17-24
A	EP,A,0 370 458 (ABBOT LABORATORIES) 30 Mai 1990 * Document entier *	1, 10, 17-19, 21, 23
Y	EP,A,0 245 136 (TRANSGENE S.A. ET AL.) 11 Novembre 1987 cité dans la demande * Document entier *	1-3, 10, 17-24

-/-		
<div style="display: flex; justify-content: space-between;"> <div style="width: 45%;"> <p>⁹ Catégories spéciales de documents cités:¹¹</p> <p>"A" document définissant l'état général de la technique, non considéré comme particulièrement pertinent</p> <p>"E" document antérieur, mais publié à la date de dépôt international ou après cette date</p> <p>"L" document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison spéciale (telle qu'indiquée)</p> <p>"O" document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens</p> <p>"P" document publié avant la date de dépôt international, mais postérieurement à la date de priorité revendiquée</p> </div> <div style="width: 45%;"> <p>"T" document ultérieur publié postérieurement à la date de dépôt international ou à la date de priorité et n'appartenant pas à l'état de la technique pertinent, mais cité pour comprendre le principe ou la théorie constituant la base de l'invention</p> <p>"X" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive</p> <p>"Y" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier.</p> <p>"&" document qui fait partie de la même famille de brevets</p> </div> </div>		
IV. CERTIFICATION		
Date à laquelle la recherche internationale a été effectivement achevée	Date d'expédition du présent rapport de recherche internationale	
06 AOUT 1992	07 SEP 1992	
Administration chargée de la recherche internationale	Signature du fonctionnaire autorisé	
OFFICE EUROPEEN DES BREVETS	JULIA P.	

III. DOCUMENTS CONSIDERES COMME PERTINENTS ¹⁴		(SUITE DES RENSEIGNEMENTS INDIQUEES SUR LA DEUXIEME FEUILLE)
Catégorie *	Identification des documents cités, ¹⁶ avec indication, si nécessaire des passages pertinents ¹⁷	No. des revendications visées ¹⁸
Y	WO,A,9 012 880 (APPLIED BIOTECHNOLOGY, INC.) 1 Novembre 1990 * Document entier *	1-3,10, 17-24
P,X	WO,A,9 105 864 (CONNAUGHT LABORATORIES LIMITED) 2 Mai 1991 * Document entier *	1-6, 10-14, 17-19, 21,23
P,X	WO,A,9 107 664 (CAMBRIDGE BIOSCIENCE CORPORATION) 30 Mai 1991 * Document entier *	1,10, 17-19, 21,23
P,X	WO,A,9 115 238 (GENENTECH, INC.) 17 Octobre 1991 * Document entier *	1,10, 17-19, 21,23

**ANNEXE AU RAPPORT DE RECHERCHE INTERNATIONALE
RELATIF A LA DEMANDE INTERNATIONALE NO.**

FR 9200394
SA 59743

La présente annexe indique les membres de la famille de brevets relatifs aux documents brevets cités dans le rapport de recherche internationale visé ci-dessus.

Lesdits membres sont contenus au fichier informatique de l'Office européen des brevets à la date du

Les renseignements fournis sont donnés à titre indicatif et n'engagent pas la responsabilité de l'Office européen des brevets. 06/08/92

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)	Date de publication	
EP-A-0314534	03-05-89	FR-A-	2620030	10-03-89
		AU-A-	2190288	09-03-89
		JP-A-	2000448	05-01-90

WO-A-9102544	07-03-91	FR-A-	2650954	22-02-91
		EP-A-	0439601	07-08-91

EP-A-0370458	30-05-90	AU-A-	4545889	05-07-90
		CA-A-	2003383	23-05-90
		JP-A-	2273187	07-11-90

EP-A-0245136	11-11-87	FR-A-	2596771	09-10-87
		FR-A,B	2606029	06-05-88
		AU-B-	604696	03-01-91
		AU-A-	7234987	09-11-87
		WO-A-	8706260	22-10-87
		JP-T-	1500161	26-01-89

WO-A-9012880	01-11-90	EP-A-	0469089	05-02-92

WO-A-9105864	02-05-91	EP-A-	0495811	29-07-92

WO-A-9107664	30-05-91	EP-A-	0453563	30-10-91

WO-A-9115238	17-10-91	AU-A-	7584791	30-10-91

EPO FORM P0472

Pour tout renseignement concernant cette annexe : voir Journal Officiel de l'Office européen des brevets, No.12/82